

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 10:14:34 ; Search time 300 Seconds  
(without alignments)  
10022.530 Million cell updates/sec

Title: US-10-033-190-1

Perfect score: 825  
Sequence: 1 atgaacagcatcatctatgc.....tatgaatctacttgatctaa 825

Scoring table: OLIGO\_NTC  
Gapop 60.0 , Gapext 60.0

Searched: 2434939 segs, 1822278265 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/US05\_PUBCOMB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/US04\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/US03\_PUBCOMB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/US02\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/US01\_PUBCOMB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/US00\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 12: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 13: /cgn2\_6/prodata/1/pubpna/US05\_PUBCOMB.seq:\*
- 14: /cgn2\_6/prodata/1/pubpna/US04\_PUBCOMB.seq:\*
- 15: /cgn2\_6/prodata/1/pubpna/US03\_PUBCOMB.seq:\*
- 16: /cgn2\_6/prodata/1/pubpna/US02\_PUBCOMB.seq:\*
- 17: /cgn2\_6/prodata/1/pubpna/US01\_PUBCOMB.seq:\*
- 18: /cgn2\_6/prodata/1/pubpna/US00\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	825	100.0	US-10-033-190-1	Sequence 1, Appl
2	567	68.7	US-10-033-190-4	Sequence 45, Appl
3	23	2.8	US-10-021-811-45	Sequence 251, Appl
4	23	2.8	US-09-934-455-251	Sequence 43, Appl
5	23	2.8	US-10-021-811-43	Sequence 2454, Ap
6	21	2.5	US-09-878-574-244	Sequence 121, App
7	21	2.5	US-10-329-960-1	Sequence 121, App
8	20	2.4	US-09-938-842A-131	Sequence 24702, A
9	20	2.4	US-09-938-842A-131	Sequence 91990, A
10	20	2.4	US-10-369-493-24702	Sequence 31145, A
11	19	2.3	US-10-098-2633-91990	Sequence 31145, A
12	19	2.3	US-10-027-632-31145	Sequence 101162, A
13	19	2.3	US-10-027-632-31145	Sequence 101162, A
14	19	2.3	US-10-027-632-101162	Sequence 101162, A
15	19	2.3	US-10-027-632-101162	Sequence 101162, A

16	19	2.3	1137	12	US-10-302-267-151	Sequence 151, App
17	19	2.3	1394	15	US-10-114-170-30	Sequence 30, Appl
18	19	2.3	1588	12	US-10-369-493-45683	Sequence 45683, A
19	19	2.3	1911	15	US-10-076-157-1	Sequence 1, Appl
20	18	2.2	95	9	US-09-864-761-21100	Sequence 21100, A
21	18	2.2	452	9	US-09-770-444-647	Sequence 647, App
22	18	2.2	452	9	US-09-864-761-4350	Sequence 4350, Ap
23	18	2.2	468	12	US-10-242-533A-19812	Sequence 39812, A
24	18	2.2	534	11	US-10-242-533A-16017	Sequence 16017, A
25	18	2.2	541	12	US-09-991-938-1826	Sequence 1826, Ap
26	18	2.2	559	13	US-10-027-632-193231	Sequence 193231, A
27	18	2.2	559	13	US-10-027-632-193232	Sequence 193232, A
28	18	2.2	559	13	US-10-027-632-193231	Sequence 193231, A
29	18	2.2	559	13	US-10-027-632-193232	Sequence 193232, A
30	18	2.2	579	14	US-10-027-632-275468	Sequence 275468, A
31	18	2.2	579	14	US-10-027-632-275468	Sequence 275468, A
32	18	2.2	586	13	US-10-027-632-230601	Sequence 230601, A
33	18	2.2	586	13	US-10-027-632-230602	Sequence 230602, A
34	18	2.2	586	14	US-10-027-632-230601	Sequence 230601, A
35	18	2.2	586	14	US-10-027-632-230602	Sequence 230602, A
36	18	2.2	612	9	US-09-770-149-860	Sequence 860, App
37	18	2.2	612	11	US-09-934-455-389	Sequence 389, App
38	18	2.2	612	12	US-10-407-920-28	Sequence 28, Appl
39	18	2.2	648	13	US-10-027-632-284815	Sequence 284815, A
40	18	2.2	648	14	US-10-027-632-284815	Sequence 284815, A
41	18	2.2	674	13	US-10-027-632-257383	Sequence 257383, A
42	18	2.2	674	14	US-10-027-632-257383	Sequence 257383, A
43	18	2.2	750	10	US-09-938-842A-331	Sequence 331, App
44	18	2.2	750	12	US-09-938-842A-331	Sequence 331, App
45	18	2.2	770	13	US-10-027-632-118985	Sequence 118985, A
46	18	2.2	770	14	US-10-027-632-118985	Sequence 118985, A
47	18	2.2	982	9	US-09-770-445-268	Sequence 268, App
48	18	2.2	984	15	US-10-278-177-21	Sequence 21, Appl
49	18	2.2	1033	11	US-10-295-403-43	Sequence 43, Appl
50	18	2.2	1033	11	US-09-934-455-13	Sequence 13, Appl
51	18	2.2	1050	15	US-10-278-177-137	Sequence 137, App
52	18	2.2	1050	16	US-10-278-536-113	Sequence 113, App
53	18	2.2	1141	9	US-09-443-704-11	Sequence 11, Appl
54	18	2.2	1141	14	US-10-008-118A-11	Sequence 11, Appl
55	18	2.2	1141	15	US-10-295-403-65	Sequence 65, Appl
56	18	2.2	1151	16	US-10-278-536-67	Sequence 67, Appl
57	18	2.2	1322	11	US-09-934-455-423	Sequence 423, App
58	18	2.2	1450	12	US-10-369-493-26114	Sequence 26114, A
59	18	2.2	1600	12	US-10-302-267-91	Sequence 91, Appl
60	18	2.2	2289	13	US-10-369-493-36473	Sequence 36473, A
61	18	2.2	2787	12	US-10-032-585-6576	Sequence 6576, App
62	18	2.2	3335	13	US-09-946-290-21	Sequence 21, Appl
63	18	2.2	3757	11	US-09-892-398-28	Sequence 28, Appl
64	18	2.2	3757	11	US-09-892-398-30	Sequence 30, Appl
65	18	2.2	9513	9	US-09-764-669-1811	Sequence 1811, App
66	18	2.2	9513	12	US-10-227-577-1811	Sequence 1811, App
67	18	2.2	9513	15	US-10-091-504-1811	Sequence 1811, App
68	18	2.2	11427	9	US-09-764-869-1812	Sequence 1812, App
69	18	2.2	11427	12	US-10-227-577-1812	Sequence 1812, App
70	18	2.2	11427	15	US-10-091-504-1812	Sequence 1812, App
71	18	2.2	15951	13	US-10-311-455-1654	Sequence 1654, App
72	18	2.2	15951	14	US-10-240-485-1317	Sequence 1317, A
73	18	2.2	17674	13	US-10-311-455-1318	Sequence 1318, App
74	18	2.2	32131	10	US-09-764-869-1678	Sequence 1678, App
75	18	2.2	32131	11	US-09-901-136-3	Sequence 3, Appl
76	18	2.2	715517	13	US-10-027-632-53712	Sequence 53712, A
77	18	2.2	715517	14	US-10-027-632-53712	Sequence 53712, A
78	18	2.2	715517	14	US-10-242-533A-13177	Sequence 13177, A
79	17	2.1	154	12	US-10-027-632-292286	Sequence 292286, A
80	17	2.1	171	13	US-10-027-632-292286	Sequence 292286, A
81	17	2.1	171	14	US-10-027-632-292286	Sequence 292286, A
82	17	2.1	201	13	US-10-168-038-26613	Sequence 26613, A
83	17	2.1	201	13	US-10-168-038-26613	Sequence 26613, A
84	17	2.1	232	9	US-09-864-761-27068	Sequence 27068, A
85	17	2.1	232	9	US-09-867-550-509	Sequence 509, App
86	17	2.1	282	10	US-09-878-574-12778	Sequence 12778, A
87	17	2.1	303	13	US-10-168-097A-47	Sequence 47, Appl
88	17	2.1	330	12	US-09-864-408A-4531	Sequence 4531, App

c 89	17	2.1	411	11	US-09-918-995-6046	Sequence 6046, Ap	c 162	17	2.1	804	13	US-10-027-632-28939	Sequence 28939, A
c 90	17	2.1	413	13	US-09-983-965-3166	Sequence 3166, Ap	c 163	17	2.1	808	14	US-10-027-632-28939	Sequence 28939, A
c 91	17	2.1	430	13	US-10-027-632-80471	Sequence 80471, A	c 164	17	2.1	804	13	US-10-027-632-28939	Sequence 28939, A
c 92	17	2.1	430	13	US-10-027-632-315111	Sequence 315111, A	c 165	17	2.1	808	14	US-10-027-632-28939	Sequence 28939, A
c 93	17	2.1	430	14	US-10-027-632-80471	Sequence 80471, A	c 166	17	2.1	833	13	US-10-027-632-16689	Sequence 16689, A
c 94	17	2.1	430	14	US-10-027-632-315111	Sequence 315111, A	c 167	17	2.1	833	13	US-10-027-632-16689	Sequence 16689, A
c 95	17	2.1	445	13	US-10-027-632-174551	Sequence 174551, A	c 168	17	2.1	833	14	US-10-027-632-16689	Sequence 16689, A
c 96	17	2.1	445	13	US-10-027-632-299035	Sequence 299035, A	c 169	17	2.1	833	14	US-10-027-632-16689	Sequence 16689, A
c 97	17	2.1	445	14	US-10-027-632-174551	Sequence 174551, A	c 170	17	2.1	838	13	US-10-027-632-156779	Sequence 156779, A
c 98	17	2.1	445	14	US-10-027-632-299035	Sequence 299035, A	c 171	17	2.1	838	13	US-10-027-632-156779	Sequence 156779, A
c 99	17	2.1	455	13	US-10-027-632-202246	Sequence 202246, A	c 172	17	2.1	838	13	US-10-027-632-156779	Sequence 156779, A
c 100	17	2.1	455	14	US-10-027-632-202246	Sequence 202246, A	c 173	17	2.1	838	14	US-10-027-632-156779	Sequence 156779, A
c 101	17	2.1	455	9	US-09-864-761-10439	Sequence 10439, A	c 174	17	2.1	838	14	US-10-027-632-156779	Sequence 156779, A
c 102	17	2.1	469	11	US-09-918-995-13314	Sequence 13314, A	c 175	17	2.1	838	14	US-10-027-632-156779	Sequence 156779, A
c 103	17	2.1	475	11	US-09-918-995-5413	Sequence 5413, Ap	c 176	17	2.1	857	14	US-10-027-632-156779	Sequence 156779, A
c 104	17	2.1	486	11	US-09-918-995-4771	Sequence 4771, Ap	c 177	17	2.1	857	14	US-10-027-632-156779	Sequence 156779, A
c 105	17	2.1	500	12	US-10-242-535A-52007	Sequence 52007, A	c 178	17	2.1	871	13	US-10-027-632-252781	Sequence 252781, A
c 106	17	2.1	514	9	US-09-443-704-5	Sequence 5, Appl1	c 179	17	2.1	871	13	US-10-027-632-252781	Sequence 252781, A
c 107	17	2.1	514	14	US-10-008-118A-5	Sequence 5, Appl1	c 180	17	2.1	871	13	US-10-027-632-252781	Sequence 252781, A
c 108	17	2.1	522	13	US-10-027-632-283125	Sequence 283125, A	c 181	17	2.1	871	14	US-10-027-632-252781	Sequence 252781, A
c 109	17	2.1	522	14	US-10-027-632-283125	Sequence 283125, A	c 182	17	2.1	871	14	US-10-027-632-252781	Sequence 252781, A
c 110	17	2.1	525	13	US-10-027-632-81622	Sequence 81622, A	c 183	17	2.1	871	14	US-10-027-632-252781	Sequence 252781, A
c 111	17	2.1	525	14	US-10-027-632-81622	Sequence 81622, A	c 184	17	2.1	888	9	US-09-770-445-502	Sequence 502, App
c 112	17	2.1	539	13	US-10-027-632-38200	Sequence 38200, A	c 185	17	2.1	888	10	US-09-938-842A-1590	Sequence 1590, Ap
c 113	17	2.1	539	14	US-10-027-632-38200	Sequence 38200, A	c 186	17	2.1	888	12	US-09-938-842A-1590	Sequence 1590, Ap
c 114	17	2.1	543	12	US-10-260-238-4970	Sequence 4970, Ap	c 187	17	2.1	897	9	US-09-864-675-3	Sequence 3, Appl1
c 115	17	2.1	552	13	US-10-027-632-199210	Sequence 199210, Ap	c 188	17	2.1	897	10	US-09-938-842A-984	Sequence 984, App
c 116	17	2.1	552	14	US-10-027-632-199210	Sequence 199210, Ap	c 189	17	2.1	897	12	US-09-938-842A-984	Sequence 984, App
c 117	17	2.1	558	13	US-10-027-632-210196	Sequence 210196, A	c 190	17	2.1	933	10	US-09-938-842A-2243	Sequence 2243, Ap
c 118	17	2.1	558	14	US-10-027-632-210196	Sequence 210196, A	c 191	17	2.1	933	12	US-09-938-842A-2243	Sequence 2243, Ap
c 119	17	2.1	562	13	US-10-027-632-289726	Sequence 289726, A	c 192	17	2.1	960	10	US-09-938-842A-2138	Sequence 2138, Ap
c 120	17	2.1	562	14	US-10-027-632-289726	Sequence 289726, A	c 193	17	2.1	960	12	US-09-938-842A-2138	Sequence 2138, Ap
c 121	17	2.1	570	13	US-10-027-632-197654	Sequence 197654, A	c 194	17	2.1	974	9	US-09-934-455-293	Sequence 293, App
c 122	17	2.1	570	13	US-10-027-632-197654	Sequence 197654, A	c 195	17	2.1	974	10	US-09-770-445-281	Sequence 281, App
c 123	17	2.1	570	14	US-10-027-632-197654	Sequence 197654, A	c 196	17	2.1	978	10	US-09-764-864-748	Sequence 748, App
c 124	17	2.1	570	14	US-10-027-632-197654	Sequence 197654, A	c 197	17	2.1	990	10	US-09-895-912A-121	Sequence 121, App
c 125	17	2.1	570	14	US-10-027-632-197655	Sequence 197655, A	c 198	17	2.1	990	9	US-09-864-675-1	Sequence 1, Appl1
c 126	17	2.1	573	13	US-10-027-632-197656	Sequence 197656, A	c 199	17	2.1	1010	12	US-10-286-264-153	Sequence 163, App
c 127	17	2.1	573	13	US-10-029-386-12923	Sequence 12923, A	c 200	17	2.1	1011	15	US-10-302-267-145	Sequence 145, App
c 128	17	2.1	582	10	US-09-796-692-3960	Sequence 3960, Ap	c 201	17	2.1	1011	15	US-10-302-267-145	Sequence 145, App
c 129	17	2.1	582	12	US-10-057-475B-3960	Sequence 3960, Ap	c 202	17	2.1	1043	13	US-10-027-632-118076	Sequence 118076, A
c 130	17	2.1	582	12	US-10-154-884B-3960	Sequence 3960, Ap	c 203	17	2.1	1043	13	US-10-027-632-118077	Sequence 118077, A
c 131	17	2.1	582	15	US-10-040-862-3960	Sequence 3960, Ap	c 204	17	2.1	1043	14	US-10-027-632-118076	Sequence 118076, A
c 132	17	2.1	595	13	US-10-027-632-81634	Sequence 81634, A	c 205	17	2.1	1043	14	US-10-027-632-118077	Sequence 118077, A
c 133	17	2.1	595	14	US-10-027-632-81634	Sequence 81634, A	c 206	17	2.1	1059	10	US-09-938-842A-1488	Sequence 1488, Ap
c 134	17	2.1	609	13	US-10-027-632-108661	Sequence 108661, A	c 207	17	2.1	1059	12	US-09-938-842A-1488	Sequence 1488, Ap
c 135	17	2.1	609	13	US-10-027-632-312709	Sequence 312709, A	c 208	17	2.1	1091	11	US-09-822-846-89	Sequence 89, Appl1
c 136	17	2.1	609	14	US-10-027-632-108661	Sequence 108661, A	c 209	17	2.1	1101	12	US-10-369-493-34050	Sequence 34050, A
c 137	17	2.1	609	14	US-10-027-632-312709	Sequence 312709, A	c 210	17	2.1	1101	12	US-10-369-493-34050	Sequence 34050, A
c 138	17	2.1	620	10	US-09-881-752A-63	Sequence 63, Appl1	c 211	17	2.1	1115	13	US-09-938-842A-2333	Sequence 2333, Ap
c 139	17	2.1	621	13	US-10-027-632-256671	Sequence 256671, A	c 212	17	2.1	1115	13	US-09-938-842A-2333	Sequence 2333, Ap
c 140	17	2.1	621	14	US-10-027-632-256671	Sequence 256671, A	c 213	17	2.1	1115	13	US-10-027-632-117277	Sequence 117277, A
c 141	17	2.1	633	13	US-10-027-632-204979	Sequence 204979, A	c 214	17	2.1	1115	13	US-10-027-632-117277	Sequence 117277, A
c 142	17	2.1	633	14	US-10-027-632-204979	Sequence 204979, A	c 215	17	2.1	1115	14	US-10-027-632-117278	Sequence 117278, A
c 143	17	2.1	648	13	US-10-027-632-224168	Sequence 224168, A	c 216	17	2.1	1115	16	US-10-027-632-117278	Sequence 117278, A
c 144	17	2.1	648	14	US-10-027-632-224168	Sequence 224168, A	c 217	17	2.1	1115	16	US-10-027-632-117278	Sequence 117278, A
c 145	17	2.1	659	13	US-10-027-632-25250	Sequence 25250, A	c 218	17	2.1	1115	12	US-10-242-355-976	Sequence 976, App
c 146	17	2.1	659	14	US-10-027-632-25250	Sequence 25250, A	c 219	17	2.1	1115	12	US-10-242-355-976	Sequence 976, App
c 147	17	2.1	660	11	US-09-782-974C-47	Sequence 47, Appl1	c 220	17	2.1	1115	12	US-10-242-355-976	Sequence 976, App
c 148	17	2.1	663	13	US-10-027-632-237299	Sequence 237299, A	c 221	17	2.1	1115	12	US-10-242-355-976	Sequence 976, App
c 149	17	2.1	663	14	US-10-027-632-237299	Sequence 237299, A	c 222	17	2.1	1115	12	US-10-242-355-976	Sequence 976, App
c 150	17	2.1	677	9	US-09-828-644-46	Sequence 46, Appl1	c 223	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1
c 151	17	2.1	708	13	US-10-027-632-16262	Sequence 16262, A	c 224	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1
c 152	17	2.1	708	13	US-10-027-632-16262	Sequence 16262, A	c 225	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1
c 153	17	2.1	708	13	US-10-027-632-16262	Sequence 16262, A	c 226	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1
c 154	17	2.1	708	14	US-10-027-632-16262	Sequence 16262, A	c 227	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1
c 155	17	2.1	708	14	US-10-027-632-16262	Sequence 16262, A	c 228	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1
c 156	17	2.1	708	14	US-10-027-632-16262	Sequence 16262, A	c 229	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1
c 157	17	2.1	714	13	US-10-027-632-150333	Sequence 150333, A	c 230	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1
c 158	17	2.1	714	14	US-10-027-632-150333	Sequence 150333, A	c 231	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1
c 159	17	2.1	747	12	US-10-369-493-37005	Sequence 37005, A	c 232	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1
c 160	17	2.1	791	9	US-09-770-445-845	Sequence 845, App	c 233	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1
c 161	17	2.1	795	9	US-09-770-445-833	Sequence 833, App	c 234	17	2.1	1236	14	US-10-050-786-4	Sequence 4, Appl1



235	17	2.1	1476	14	US-10-096-241-7	Sequence 7, Appli	308	17	2.1	744802	12	US-10-292-798-1369	Sequence 1369, Ap
236	17	2.1	1504	9	US-09-925-301-258	Sequence 258, App	309	17	2.1	786430	13	US-10-412-277-3	Sequence 3, Appli
C 237	17	2.1	1527	10	US-09-974-298-171	Sequence 171, App	310	17	2.1	1503841	9	US-09-795-668-1	Sequence 1, Appli
C 238	17	2.1	1537	10	US-09-981-353-139	Sequence 139, App	311	17	2.1	1503841	9	US-09-795-668-1	Sequence 1, Appli
239	17	2.1	1552	13	US-10-027-632-263186	Sequence 263186,	312	17	2.1	1503841	10	US-09-946-807-1	Sequence 1, Appli
240	17	2.1	1552	14	US-10-027-632-263186	Sequence 263186,	313	17	2.1	1503841	10	US-09-946-807-1	Sequence 1, Appli
241	17	2.1	1550	13	US-10-342-224-111	Sequence 111, App	C 314	17	2.1	2140405	13	US-10-027-632-76212	Sequence 76212, A
242	17	2.1	1576	10	US-09-938-842A-3682	Sequence 3682, Ap	C 315	17	2.1	2140405	14	US-10-027-632-76212	Sequence 76212, A
243	17	2.1	1576	12	US-09-938-842A-3682	Sequence 3682, Ap	C 316	17	2.1	3673778	13	US-10-312-841-1	Sequence 1, Appli
244	17	2.1	1588	13	US-10-168-097A-45	Sequence 45, Appl	C 317	17	2.1	3673778	13	US-10-312-841-2	Sequence 2, Appli
C 245	17	2.1	1604	13	US-10-085-959-188	Sequence 188, App	318	16	1.9	25	15	US-10-098-263B-4673	Sequence 64673, A
C 246	17	2.1	1629	13	US-10-158-646-13	Sequence 13, Appl	319	16	1.9	48	11	US-09-976-800-68	Sequence 68, Appl
C 247	17	2.1	1722	12	US-10-138-646-13	Sequence 13, Appl	320	16	1.9	48	13	US-10-139-296-68	Sequence 68, Appl
C 248	17	2.1	1766	9	US-09-817-910-1	Sequence 1, Appli	321	16	1.9	48	13	US-10-139-296-68	Sequence 68, Appl
249	17	2.1	1797	12	US-10-369-493-43475	Sequence 43475, A	322	16	1.9	48	13	US-10-405-660-68	Sequence 68, Appl
C 250	17	2.1	1884	14	US-10-096-241-5	Sequence 5, Appli	323	16	1.9	48	15	US-10-138-839-68	Sequence 68, Appl
C 251	17	2.1	1937	13	US-10-108-260A-219	Sequence 219, App	324	16	1.9	48	15	US-10-138-839-68	Sequence 68, Appl
C 252	17	2.1	1939	13	US-10-027-632-97312	Sequence 97312, A	325	16	1.9	48	15	US-10-138-839-68	Sequence 68, Appl
253	17	2.1	1999	14	US-10-027-632-97312	Sequence 97312, A	326	16	1.9	48	15	US-10-138-839-68	Sequence 68, Appl
254	17	2.1	2000	10	US-09-938-842A-4414	Sequence 4414, Ap	327	16	1.9	80	13	US-10-198-235-3	Sequence 3, Appli
255	17	2.1	2000	12	US-09-938-842A-4414	Sequence 4414, Ap	C 328	16	1.9	103	12	US-10-242-535A-12880	Sequence 32880, A
256	17	2.1	2015	10	US-09-764-864-331	Sequence 331, App	C 329	16	1.9	147	10	US-09-960-352-14450	Sequence 14450, A
257	17	2.1	2110	9	US-09-938-803-36	Sequence 36, Appl	C 330	16	1.9	160	9	US-09-864-761-17841	Sequence 17841, A
258	17	2.1	2142	10	US-09-917-800A-1567	Sequence 1567, Ap	331	16	1.9	168	10	US-09-974-300-217	Sequence 217, App
259	17	2.1	2142	12	US-10-191-803-35	Sequence 35, Appl	332	16	1.9	208	9	US-09-815-242-1395	Sequence 1395, Ap
C 260	17	2.1	2215	12	US-10-104-047-105	Sequence 105, App	C 333	16	1.9	217	12	US-10-242-535A-880	Sequence 880, App
261	17	2.1	2231	13	US-10-027-632-103598	Sequence 103598,	334	16	1.9	219	13	US-10-168-097A-21	Sequence 21, Appl
262	17	2.1	2231	14	US-10-027-632-103598	Sequence 103598,	335	16	1.9	226	10	US-09-983-965-5653	Sequence 5653, Ap
263	17	2.1	2268	14	US-10-096-241-31	Sequence 31, Appl	336	16	1.9	240	12	US-10-242-535A-51117	Sequence 51117, Ap
264	17	2.1	2269	12	US-10-108-260A-51	Sequence 51, Appl	C 337	16	1.9	240	12	US-10-242-535A-51117	Sequence 51117, Ap
265	17	2.1	2275	12	US-10-104-047-486	Sequence 486, App	338	16	1.9	241	12	US-10-242-535A-12250	Sequence 49944, A
C 266	17	2.1	2303	13	US-10-094-749-547	Sequence 547, App	C 339	16	1.9	249	10	US-09-878-574-9068	Sequence 9068, Ap
267	17	2.1	2391	13	US-10-032-585-6201	Sequence 6201, Ap	340	16	1.9	255	14	US-10-047-412A-26	Sequence 26, Appl
268	17	2.1	2408	12	US-10-104-047-1763	Sequence 1763, Ap	C 341	16	1.9	257	12	US-10-242-535A-35005	Sequence 35005, A
269	17	2.1	2639	12	US-10-104-047-684	Sequence 684, App	342	16	1.9	267	13	US-10-029-386-19763	Sequence 19763, A
C 270	17	2.1	2647	13	US-10-288-252-23	Sequence 23, Appl	343	16	1.9	279	12	US-09-864-408A-4871	Sequence 4871, Ap
271	17	2.1	2834	11	US-09-845-020A-6	Sequence 6, Appl	344	16	1.9	288	13	US-10-259-165-703	Sequence 703, App
272	17	2.1	3178	11	US-09-308-207-7	Sequence 7, Appli	345	16	1.9	292	10	US-09-736-457-1781	Sequence 1781, Ap
273	17	2.1	3214	13	US-10-115-831-110	Sequence 110, Appl	346	16	1.9	292	10	US-09-902-941-1781	Sequence 1781, Ap
274	17	2.1	3465	15	US-10-224-624-9	Sequence 9, Appli	347	16	1.9	292	10	US-09-849-628-1781	Sequence 1781, Ap
275	17	2.1	3832	15	US-10-224-624-7	Sequence 7, Appli	348	16	1.9	292	15	US-10-113-872-1781	Sequence 1781, Ap
276	17	2.1	4098	12	US-10-264-049-377	Sequence 377, App	349	16	1.9	292	15	US-10-017-754-1781	Sequence 1781, Ap
277	17	2.1	4514	11	US-09-252-088-1	Sequence 1, Appli	350	16	1.9	303	13	US-10-168-097A-17	Sequence 17, Appl
278	17	2.1	4514	12	US-10-340-792-1	Sequence 1, Appli	351	16	1.9	327	12	US-10-242-535A-7046	Sequence 7046, Ap
279	17	2.1	4775	8	US-08-781-986A-203	Sequence 203, App	352	16	1.9	327	15	US-10-066-543-2889	Sequence 2889, Ap
C 280	17	2.1	5470	10	US-09-764-864-1533	Sequence 1533, Ap	353	16	1.9	332	10	US-09-867-701-4264	Sequence 4264, Ap
281	17	2.1	6235	11	US-09-845-020A-5	Sequence 5, Appli	354	16	1.9	333	10	US-09-796-692-5294	Sequence 5294, Ap
C 282	17	2.1	6265	9	US-09-129-112-3	Sequence 3, Appli	355	16	1.9	333	12	US-10-057-475B-5294	Sequence 5294, Ap
283	17	2.1	6577	13	US-09-880-107-2187	Sequence 2187, Ap	356	16	1.9	333	12	US-10-154-884B-5294	Sequence 5294, Ap
284	17	2.1	6597	13	US-10-210-120-7	Sequence 7, Appli	357	16	1.9	333	15	US-10-040-867-5294	Sequence 5294, Ap
285	17	2.1	6679	15	US-09-845-020A-1	Sequence 1, Appli	C 358	16	1.9	334	10	US-09-981-353-149	Sequence 149, App
C 286	17	2.1	6944	15	US-10-172-086-112	Sequence 112, App	C 359	16	1.9	334	15	US-10-158-646-43	Sequence 43, Appl
C 287	17	2.1	6971	13	US-10-311-455-1210	Sequence 1210, Ap	C 360	16	1.9	335	13	US-10-027-632-28301	Sequence 28301, A
C 288	17	2.1	7110	13	US-10-311-455-1529	Sequence 1529, Ap	C 361	16	1.9	335	13	US-10-027-632-28302	Sequence 28302, A
C 289	17	2.1	11968	11	US-09-764-891-8251	Sequence 8251, Ap	C 362	16	1.9	335	14	US-10-027-632-28301	Sequence 28301, A
C 290	17	2.1	13582	12	US-10-292-798-11235	Sequence 1325, Ap	C 363	16	1.9	335	14	US-10-027-632-28302	Sequence 28302, A
C 291	17	2.1	13582	13	US-10-017-161-1655	Sequence 1665, Ap	C 364	16	1.9	342	13	US-10-168-097A-13	Sequence 13, Appl
292	17	2.1	15881	13	US-10-311-455-233	Sequence 233, App	C 365	16	1.9	347	15	US-10-313-542-188	Sequence 188, App
293	17	2.1	16373	13	US-10-311-455-592	Sequence 592, App	C 366	16	1.9	355	13	US-09-814-353-12989	Sequence 12989, A
C 294	17	2.1	17421	15	US-10-240-453-55	Sequence 55, Appl	367	16	1.9	359	10	US-09-878-574-3460	Sequence 3460, App
C 295	17	2.1	17421	15	US-10-239-676-53	Sequence 53, Appl	368	16	1.9	360	11	US-09-991-936-427	Sequence 427, App
C 296	17	2.1	20530	11	US-09-764-891-88252	Sequence 8252, Ap	C 369	16	1.9	367	10	US-09-878-574-1861	Sequence 1861, Ap
297	17	2.1	20530	12	US-10-074-024-867	Sequence 867, App	C 370	16	1.9	376	11	US-09-803-719-1395	Sequence 1395, Ap
298	17	2.1	54000	13	US-09-843-377-11	Sequence 11, Appl	371	16	1.9	388	11	US-09-803-719-1395	Sequence 1395, Ap
C 299	17	2.1	58448	13	US-10-017-161-1659	Sequence 1659, Ap	372	16	1.9	396	13	US-10-027-632-28301	Sequence 28301, A
C 300	17	2.1	136726	12	US-10-085-117-244	Sequence 244, App	373	16	1.9	396	13	US-10-027-632-28301	Sequence 28301, A
301	17	2.1	177249	13	US-10-085-117-223	Sequence 223, App	374	16	1.9	396	14	US-10-027-632-28301	Sequence 28301, A
302	17	2.1	240825	12	US-09-790-289-1	Sequence 1, Appli	375	16	1.9	396	14	US-10-027-632-28301	Sequence 28301, A
C 303	17	2.1	271990	12	US-10-345-072-87	Sequence 87, Appl	376	16	1.9	398	10	US-09-764-877-404	Sequence 404, App
C 304	17	2.1	271990	15	US-10-195-144-87	Sequence 87, Appl	377	16	1.9	398	12	US-10-242-515-404	Sequence 404, App
C 305	17	2.1	513509	11	US-09-754-853A-4	Sequence 4, Appli	378	16	1.9	400	12	US-10-191-803-428	Sequence 428, App
306	17	2.1	640681	10	US-09-790-988-1	Sequence 1, Appli	379	16	1.9	401	13	US-10-027-632-64198	Sequence 64198, A
C 307	17	2.1	640681	10	US-09-790-988-1	Sequence 1, Appli	380	16	1.9	401	14	US-10-027-632-64198	Sequence 64198, A

C 381	16	1.9	408	10	US-09-938-842A-1448	Sequence 1448, Ap	C 454	16	1.9	514	13	US-09-738-269-14	Sequence 14, Appl
C 382	16	1.9	408	12	US-09-938-842A-1448	Sequence 1448, Ap	C 455	16	1.9	514	13	US-10-027-633-297666	Sequence 297666,
C 383	16	1.9	410	10	US-09-960-352-13656	Sequence 13656, Ap	C 456	16	1.9	514	13	US-10-027-633-334358	Sequence 334358,
C 384	16	1.9	410	10	US-09-933-797-199	Sequence 199, App	C 457	16	1.9	514	14	US-10-023-437-14	Sequence 14, Appl
C 385	16	1.9	412	13	US-10-101-510-703	Sequence 703, App	C 458	16	1.9	514	14	US-10-027-632-227666	Sequence 227666,
C 386	16	1.9	429	13	US-10-027-632-1254	Sequence 1254, Ap	C 459	16	1.9	514	14	US-10-027-633-334358	Sequence 334358,
C 387	16	1.9	429	14	US-10-027-632-1254	Sequence 1254, Ap	C 460	16	1.9	515	11	US-09-814-353-1731	Sequence 1731, A
C 388	16	1.9	432	10	US-09-996-692-4122	Sequence 4122, Ap	C 461	16	1.9	525	11	US-09-918-999-5530	Sequence 5530, Ap
C 389	16	1.9	432	12	US-10-057-4758-4122	Sequence 4122, Ap	C 462	16	1.9	526	13	US-10-027-633-277211	Sequence 277211,
C 390	16	1.9	432	12	US-10-154-884B-4122	Sequence 4122, Ap	C 463	16	1.9	526	14	US-10-027-632-277211	Sequence 277211,
C 391	16	1.9	432	15	US-10-040-862-4122	Sequence 4122, Ap	C 464	16	1.9	531	11	US-09-918-999-25332	Sequence 25332, A
C 392	16	1.9	434	10	US-09-924-035A-804	Sequence 804, App	C 465	16	1.9	532	15	US-10-198-846-8566	Sequence 8566, Ap
C 393	16	1.9	435	10	US-09-796-692-8915	Sequence 8915, Ap	C 466	16	1.9	535	13	US-10-027-633-34419	Sequence 34419, A
C 394	16	1.9	435	12	US-10-057-4758-8915	Sequence 8915, Ap	C 467	16	1.9	535	14	US-10-027-633-34419	Sequence 34419, A
C 395	16	1.9	435	12	US-10-154-884B-8915	Sequence 8915, Ap	C 468	16	1.9	537	13	US-09-814-353-977	Sequence 977, App
C 396	16	1.9	435	15	US-10-040-862-8915	Sequence 8915, Ap	C 469	16	1.9	537	13	US-09-814-353-7346	Sequence 7346, Ap
C 397	16	1.9	440	13	US-10-027-632-183275	Sequence 183275,	C 470	16	1.9	539	13	US-10-027-632-230136	Sequence 230136,
C 398	16	1.9	440	14	US-10-027-632-183275	Sequence 183275,	C 471	16	1.9	539	14	US-10-027-633-230136	Sequence 230136,
C 399	16	1.9	444	13	US-10-027-632-181142	Sequence 181142,	C 472	16	1.9	541	13	US-10-027-633-43902	Sequence 43902, A
C 400	16	1.9	444	13	US-10-027-632-304689	Sequence 304689,	C 473	16	1.9	541	13	US-10-027-633-71457	Sequence 71457, A
C 401	16	1.9	444	14	US-10-027-632-181142	Sequence 181142,	C 474	16	1.9	541	13	US-10-027-633-71458	Sequence 71458, A
C 402	16	1.9	444	14	US-10-027-632-304689	Sequence 304689,	C 475	16	1.9	541	13	US-10-027-633-293760	Sequence 293760,
C 403	16	1.9	448	13	US-10-027-632-30388	Sequence 30388,	C 476	16	1.9	541	13	US-10-027-632-293761	Sequence 293761,
C 404	16	1.9	448	13	US-10-027-632-30388	Sequence 30388,	C 477	16	1.9	541	14	US-10-027-633-43902	Sequence 43902, A
C 405	16	1.9	456	13	US-10-027-632-74081	Sequence 74081, A	C 478	16	1.9	541	14	US-10-027-633-71457	Sequence 71457, A
C 406	16	1.9	456	14	US-10-027-632-74081	Sequence 74081, A	C 479	16	1.9	541	14	US-10-027-633-71458	Sequence 71458, A
C 407	16	1.9	457	11	US-09-918-999-12529	Sequence 12529, A	C 480	16	1.9	541	14	US-10-027-633-293760	Sequence 293760,
C 408	16	1.9	461	12	US-10-242-535A-36890	Sequence 36890, A	C 481	16	1.9	541	14	US-10-027-633-293761	Sequence 293761,
C 409	16	1.9	467	9	US-09-864-761-1071	Sequence 1071, App	C 482	16	1.9	542	11	US-09-991-938-1491	Sequence 1491, Ap
C 410	16	1.9	467	9	US-09-764-869-531	Sequence 531, App	C 483	16	1.9	545	13	US-10-029-386-11976	Sequence 11976, Ap
C 411	16	1.9	467	12	US-10-227-577-531	Sequence 531, App	C 484	16	1.9	547	10	US-09-728-448-70	Sequence 70, Appl
C 412	16	1.9	467	15	US-10-091-504-531	Sequence 531, App	C 485	16	1.9	547	10	US-09-728-448-414	Sequence 414, Appl
C 413	16	1.9	472	9	US-09-864-761-15020	Sequence 15020, A	C 486	16	1.9	550	13	US-10-027-633-187586	Sequence 187586,
C 414	16	1.9	472	12	US-10-242-535A-41003	Sequence 41003, A	C 487	16	1.9	550	14	US-10-027-633-187586	Sequence 187586,
C 415	16	1.9	477	13	US-10-027-632-195241	Sequence 195241,	C 488	16	1.9	551	9	US-09-820-089-13	Sequence 13, Appl
C 416	16	1.9	477	14	US-10-027-632-195241	Sequence 195241,	C 489	16	1.9	553	13	US-10-029-386-6032	Sequence 6032, Ap
C 417	16	1.9	478	13	US-10-027-632-2977	Sequence 2977, Ap	C 490	16	1.9	559	13	US-10-027-633-234835	Sequence 234835,
C 418	16	1.9	478	14	US-10-027-632-2977	Sequence 2977, Ap	C 491	16	1.9	559	14	US-10-027-633-234835	Sequence 234835,
C 419	16	1.9	479	11	US-09-918-999-23476	Sequence 23476, A	C 492	16	1.9	559	14	US-10-027-633-234835	Sequence 234835,
C 420	16	1.9	479	13	US-10-252-157-309	Sequence 309, App	C 493	16	1.9	559	14	US-10-027-633-234836	Sequence 234836,
C 421	16	1.9	479	13	US-10-027-632-88298	Sequence 88298, A	C 494	16	1.9	560	9	US-09-864-761-13746	Sequence 13746, A
C 422	16	1.9	479	14	US-10-027-632-88298	Sequence 88298, A	C 495	16	1.9	562	13	US-10-027-633-267442	Sequence 267442,
C 423	16	1.9	481	13	US-10-027-632-305910	Sequence 305910,	C 496	16	1.9	562	14	US-10-027-633-267442	Sequence 267442,
C 424	16	1.9	481	14	US-10-027-632-305910	Sequence 305910,	C 497	16	1.9	565	12	US-10-242-533A-45471	Sequence 45471, A
C 425	16	1.9	482	11	US-09-918-999-23265	Sequence 23265, A	C 498	16	1.9	569	13	US-10-027-633-88207	Sequence 88207, A
C 426	16	1.9	482	13	US-10-027-632-56313	Sequence 56313, A	C 499	16	1.9	569	13	US-10-027-633-88207	Sequence 88207, A
C 427	16	1.9	483	13	US-10-027-632-296909	Sequence 296909,	C 500	16	1.9	569	14	US-10-027-633-88207	Sequence 88207, A
C 428	16	1.9	483	14	US-10-027-632-56313	Sequence 56313, A	C 501	16	1.9	570	13	US-10-027-633-88208	Sequence 88208, A
C 429	16	1.9	483	14	US-10-027-632-296909	Sequence 296909,	C 502	16	1.9	570	14	US-10-027-633-210201	Sequence 210201,
C 430	16	1.9	485	11	US-09-918-999-6439	Sequence 6439, Ap	C 503	16	1.9	572	13	US-10-027-633-238064	Sequence 238064,
C 431	16	1.9	485	13	US-10-027-632-36616	Sequence 36616, A	C 504	16	1.9	572	13	US-10-027-633-238064	Sequence 238064,
C 432	16	1.9	485	13	US-10-027-632-69682	Sequence 69682, A	C 505	16	1.9	572	13	US-10-027-633-239015	Sequence 239015,
C 433	16	1.9	485	13	US-10-027-632-302085	Sequence 302085,	C 506	16	1.9	572	13	US-10-027-633-239016	Sequence 239016,
C 434	16	1.9	485	14	US-10-027-632-36616	Sequence 36616, A	C 507	16	1.9	572	14	US-10-027-633-239015	Sequence 239015,
C 435	16	1.9	485	14	US-10-027-632-69682	Sequence 69682, A	C 508	16	1.9	572	14	US-10-027-633-239016	Sequence 239016,
C 436	16	1.9	485	14	US-10-027-632-302085	Sequence 302085,	C 509	16	1.9	575	13	US-10-029-386-11523	Sequence 11523, A
C 437	16	1.9	489	11	US-09-918-999-24941	Sequence 24941, A	C 510	16	1.9	577	13	US-10-029-386-5187	Sequence 5187, Ap
C 438	16	1.9	493	13	US-10-027-632-94696	Sequence 94696, A	C 511	16	1.9	579	10	US-09-917-800A-709	Sequence 709, App
C 439	16	1.9	493	14	US-10-027-632-94696	Sequence 94696, A	C 512	16	1.9	580	13	US-10-027-633-215789	Sequence 215789,
C 440	16	1.9	499	13	US-10-027-632-95120	Sequence 95120, A	C 513	16	1.9	580	13	US-10-027-633-215790	Sequence 215790,
C 441	16	1.9	499	13	US-10-027-632-95121	Sequence 95121, A	C 514	16	1.9	580	14	US-10-027-633-215789	Sequence 215789,
C 442	16	1.9	499	14	US-10-027-632-95120	Sequence 95120, A	C 515	16	1.9	580	14	US-10-027-633-215790	Sequence 215790,
C 443	16	1.9	499	14	US-10-027-632-95121	Sequence 95121, A	C 516	16	1.9	580	14	US-10-027-633-215790	Sequence 215790,
C 444	16	1.9	501	11	US-09-918-999-19338	Sequence 19338, A	C 517	16	1.9	583	13	US-10-027-633-75180	Sequence 75180, A
C 445	16	1.9	501	11	US-09-918-999-19338	Sequence 19338, A	C 518	16	1.9	583	13	US-10-027-633-75181	Sequence 75181, A
C 446	16	1.9	503	11	US-09-918-999-224592	Sequence 224592,	C 519	16	1.9	583	14	US-10-027-633-75180	Sequence 75180, A
C 447	16	1.9	503	13	US-10-027-632-272592	Sequence 272592,	C 520	16	1.9	583	14	US-10-027-633-75181	Sequence 75181, A
C 448	16	1.9	503	14	US-10-027-632-272592	Sequence 272592,	C 521	16	1.9	585	13	US-10-027-633-224456	Sequence 224456,
C 449	16	1.9	503	14	US-10-027-632-312048	Sequence 312048,	C 522	16	1.9	587	14	US-10-027-633-224456	Sequence 224456,
C 450	16	1.9	504	11	US-09-770-961-350	Sequence 350, App	C 523	16	1.9	587	13	US-09-833-381-916	Sequence 916, App
C 451	16	1.9	507	11	US-10-027-632-323303	Sequence 323303,	C 524	16	1.9	587	13	US-10-027-633-239844	Sequence 239844,
C 452	16	1.9	507	14	US-10-027-632-323303	Sequence 323303,	C 525	16	1.9	587	13	US-10-027-633-239845	Sequence 239845,
C 453	16	1.9	508	13	US-10-125-968-6	Sequence 6, Appl	C 526	16	1.9	587	13	US-10-027-633-239846	Sequence 239846,

527	1.9	587	13	US-10-027-632-239847	Sequence 239847,	C 600	1.9	662	14	US-10-027-632-277507	Sequence 277507,
528	1.9	587	14	US-10-027-632-239844	Sequence 239844,	C 601	1.9	665	10	US-09-764-877-82	Sequence 82, Appl
529	1.9	587	14	US-10-027-632-239845	Sequence 239845,	C 602	1.9	665	12	US-10-242-855-82	Sequence 82,
530	1.9	587	14	US-10-027-632-239846	Sequence 239846,	C 603	1.9	670	11	US-09-764-891-6549	Sequence 6549, Ap
531	1.9	587	14	US-10-027-632-239847	Sequence 239847,	C 604	1.9	670	15	US-10-205-428-746	Sequence 746, Ap
C 532	1.9	589	9	US-09-864-761-14885	Sequence 14885, A	C 605	1.9	674	10	US-09-938-842A-4753	Sequence 4753, Ap
C 533	1.9	589	13	US-10-027-632-16842	Sequence 16842, A	C 606	1.9	676	12	US-09-938-842A-4753	Sequence 4753, Ap
C 534	1.9	589	13	US-10-027-632-16843	Sequence 16843, A	C 607	1.9	676	12	US-10-260-2328-3067	Sequence 3067, Ap
C 535	1.9	589	14	US-10-027-632-16842	Sequence 16842, A	C 608	1.9	678	13	US-10-027-632-106937	Sequence 106937,
C 536	1.9	589	14	US-10-027-632-16843	Sequence 16843, A	C 609	1.9	678	13	US-10-027-632-106938	Sequence 106938,
537	1.9	594	9	US-09-864-761-9057	Sequence 9057, Ap	C 610	1.9	678	14	US-10-027-632-106937	Sequence 106937,
538	1.9	595	9	US-09-864-761-14953	Sequence 14953, A	C 611	1.9	682	14	US-10-027-632-106938	Sequence 106938,
C 539	1.9	597	13	US-10-027-632-55011	Sequence 55011, A	C 612	1.9	682	14	US-10-027-632-133528	Sequence 133528,
C 540	1.9	597	14	US-10-027-632-313735	Sequence 313735,	C 613	1.9	682	14	US-10-027-632-133529	Sequence 133529,
C 541	1.9	597	14	US-10-027-632-55011	Sequence 55011, A	C 614	1.9	682	14	US-10-027-632-133528	Sequence 133528,
C 542	1.9	597	14	US-10-027-632-313735	Sequence 313735,	C 615	1.9	682	14	US-10-027-632-133529	Sequence 133529,
543	1.9	598	13	US-10-027-632-209057	Sequence 209057,	C 616	1.9	684	14	US-10-027-632-286277	Sequence 286277,
544	1.9	598	14	US-10-027-632-209058	Sequence 209058,	C 617	1.9	684	14	US-10-027-632-286277	Sequence 286277,
545	1.9	598	14	US-10-027-632-209057	Sequence 209057,	C 618	1.9	688	14	US-10-027-632-788	Sequence 788, App
546	1.9	598	14	US-10-027-632-209058	Sequence 209058,	C 619	1.9	688	14	US-10-027-632-788	Sequence 788, App
547	1.9	599	10	US-09-974-300-2950	Sequence 2950, Ap	C 620	1.9	689	13	US-10-027-632-138836	Sequence 138836,
548	1.9	603	12	US-10-369-493-40246	Sequence 40246, A	C 621	1.9	689	13	US-10-027-632-138836	Sequence 138836,
549	1.9	603	13	US-10-027-632-111220	Sequence 111220,	C 622	1.9	692	9	US-09-910-943-315	Sequence 315, App
550	1.9	603	13	US-10-027-632-111220	Sequence 111220,	C 623	1.9	697	13	US-10-029-386-25230	Sequence 25230, A
C 551	1.9	608	13	US-10-027-632-112154	Sequence 112154,	C 624	1.9	702	13	US-10-116-275-284	Sequence 284, App
C 552	1.9	608	13	US-10-027-632-122155	Sequence 122155,	C 625	1.9	719	13	US-10-027-632-17089	Sequence 17089, A
C 553	1.9	608	14	US-10-027-632-122154	Sequence 122154,	C 626	1.9	719	14	US-10-027-632-17089	Sequence 17089, A
C 554	1.9	608	14	US-10-027-632-122155	Sequence 122155,	C 627	1.9	720	12	US-10-260-228-5827	Sequence 5827, Ap
C 555	1.9	610	9	US-09-770-149-875	Sequence 875, App	C 628	1.9	722	14	US-10-027-632-13181	Sequence 13181, A
556	1.9	611	13	US-10-027-632-234017	Sequence 234017,	C 629	1.9	722	14	US-10-027-632-13181	Sequence 13181, A
557	1.9	611	14	US-10-027-632-234017	Sequence 234017,	C 630	1.9	732	10	US-09-938-842A-598	Sequence 598, App
558	1.9	613	13	US-10-027-632-198437	Sequence 198437,	C 631	1.9	732	12	US-09-938-842A-598	Sequence 598, App
559	1.9	613	14	US-10-027-632-198437	Sequence 198437,	C 632	1.9	732	13	US-10-027-632-103616	Sequence 103616,
560	1.9	615	13	US-10-300-072-35	Sequence 35, Appl	C 633	1.9	732	13	US-10-027-632-103616	Sequence 103616,
561	1.9	618	11	US-09-871-161-218	Sequence 218, App	C 634	1.9	738	9	US-09-764-887-581	Sequence 581, App
562	1.9	619	13	US-10-027-632-290322	Sequence 290322,	C 635	1.9	740	15	US-10-073-961-581	Sequence 581, App
563	1.9	619	14	US-10-027-632-290322	Sequence 290322,	C 636	1.9	740	13	US-10-027-632-313319	Sequence 313319,
564	1.9	624	13	US-10-027-632-321029	Sequence 321029,	C 637	1.9	740	13	US-10-027-632-313319	Sequence 313319,
565	1.9	624	13	US-10-027-632-321030	Sequence 321030,	C 638	1.9	740	14	US-10-027-632-313319	Sequence 313319,
566	1.9	624	14	US-10-027-632-321029	Sequence 321029,	C 639	1.9	740	14	US-10-027-632-313319	Sequence 313319,
567	1.9	624	14	US-10-027-632-321030	Sequence 321030,	C 640	1.9	747	10	US-09-938-842A-2634	Sequence 2634, Ap
568	1.9	625	13	US-10-027-632-238797	Sequence 238797,	C 641	1.9	753	12	US-09-938-842A-2634	Sequence 2634, Ap
569	1.9	625	14	US-10-027-632-238797	Sequence 238797,	C 642	1.9	753	14	US-10-027-632-102937	Sequence 102937,
C 570	1.9	626	14	US-10-027-632-282944	Sequence 282944,	C 643	1.9	755	13	US-10-027-632-25448	Sequence 25448, A
C 571	1.9	630	13	US-10-027-632-276896	Sequence 276896,	C 644	1.9	755	14	US-10-027-632-25448	Sequence 25448, A
C 572	1.9	630	14	US-10-027-632-276896	Sequence 276896,	C 645	1.9	755	14	US-10-027-632-25448	Sequence 25448, A
C 573	1.9	636	13	US-10-027-632-237985	Sequence 237985,	C 646	1.9	755	14	US-10-027-632-25448	Sequence 25448, A
C 574	1.9	636	14	US-10-027-632-237985	Sequence 237985,	C 647	1.9	756	13	US-10-029-386-22732	Sequence 22732, A
C 575	1.9	637	13	US-10-012-697-892	Sequence 822, App	C 648	1.9	756	13	US-10-027-632-48878	Sequence 48878, A
C 576	1.9	639	13	US-10-027-632-254007	Sequence 254007,	C 649	1.9	756	13	US-10-027-632-48878	Sequence 48878, A
C 577	1.9	639	14	US-10-027-632-254007	Sequence 254007,	C 650	1.9	756	14	US-10-027-632-48878	Sequence 48878, A
C 578	1.9	643	13	US-10-027-632-207715	Sequence 207715,	C 651	1.9	756	14	US-10-027-632-48878	Sequence 48878, A
C 579	1.9	643	14	US-10-027-632-207715	Sequence 207715,	C 652	1.9	756	14	US-10-027-632-48878	Sequence 48878, A
C 580	1.9	646	13	US-10-027-632-162413	Sequence 162413,	C 653	1.9	800	9	US-09-910-943-585	Sequence 585, App
581	1.9	646	14	US-10-027-632-162413	Sequence 162413,	C 654	1.9	803	15	US-10-286-264-129847	Sequence 129847,
582	1.9	649	13	US-10-027-632-2405	Sequence 2405, Ap	C 655	1.9	829	13	US-10-027-632-159412	Sequence 159412,
C 583	1.9	649	13	US-10-027-632-2405	Sequence 2405, Ap	C 656	1.9	829	14	US-10-027-632-159412	Sequence 159412,
C 584	1.9	649	14	US-10-027-632-2405	Sequence 2405, Ap	C 657	1.9	836	13	US-10-027-632-160348	Sequence 160348,
C 585	1.9	649	14	US-10-027-632-2405	Sequence 2405, Ap	C 658	1.9	836	13	US-10-027-632-160348	Sequence 160348,
C 586	1.9	649	14	US-10-027-632-2405	Sequence 2405, Ap	C 659	1.9	836	14	US-10-027-632-160348	Sequence 160348,
587	1.9	650	13	US-10-027-632-198728	Sequence 198728,	C 660	1.9	841	13	US-10-027-632-160348	Sequence 160348,
C 588	1.9	650	14	US-10-027-632-198728	Sequence 198728,	C 661	1.9	841	13	US-10-027-632-161288	Sequence 161288,
589	1.9	650	14	US-10-027-632-198728	Sequence 198728,	C 662	1.9	841	13	US-10-027-632-161288	Sequence 161288,
C 590	1.9	650	14	US-10-027-632-207727	Sequence 207727,	C 663	1.9	848	11	US-09-764-891-647	Sequence 647, App
C 591	1.9	653	13	US-10-027-632-55984	Sequence 55984, A	C 664	1.9	850	13	US-10-027-632-169413	Sequence 169413,
C 592	1.9	653	14	US-10-027-632-55984	Sequence 55984, A	C 665	1.9	850	13	US-10-027-632-169413	Sequence 169413,
593	1.9	655	13	US-10-027-632-284793	Sequence 284793,	C 666	1.9	850	14	US-10-027-632-169413	Sequence 169413,
594	1.9	655	14	US-10-027-632-284793	Sequence 284793,	C 667	1.9	850	14	US-10-027-632-169413	Sequence 169413,
595	1.9	657	13	US-10-027-632-278831	Sequence 278831,	C 668	1.9	855	13	US-10-027-632-169178	Sequence 169178,
596	1.9	657	13	US-10-027-632-278831	Sequence 278831,	C 669	1.9	855	14	US-10-027-632-169178	Sequence 169178,
597	1.9	657	14	US-10-027-632-278831	Sequence 278831,	C 670	1.9	874	13	US-10-259-165-623	Sequence 623, App
598	1.9	657	14	US-10-027-632-278832	Sequence 278832,	C 671	1.9	875	13	US-10-027-632-173507	Sequence 173507,
C 599	1.9	662	13	US-10-027-632-277507	Sequence 277507,	C 672	1.9	875	14	US-10-027-632-173507	Sequence 173507,

673	16	1.9	887	13	US-10-103-959A-1	Sequence 1, Appl1	C 746	16	1.9	1657	10	US-09-764-877-3597	Sequence 3597, Ap
C 674	16	1.9	924	14	US-10-027-632-10231	Sequence 10231, A	C 747	16	1.9	1657	12	US-10-242-515-3597	Sequence 3597, Ap
C 675	16	1.9	924	13	US-10-027-632-10231	Sequence 10211, A	C 748	16	1.9	1736	9	US-09-764-898-62	Sequence 62, Appl
676	16	1.9	925	11	US-09-934-455-437	Sequence 437, App	749	16	1.9	1737	10	US-09-731-897-69	Sequence 69, Appl
677	16	1.9	943	9	US-09-770-445-352	Sequence 352, App	750	16	1.9	1737	13	US-09-876-997-69	Sequence 69, Appl
678	16	1.9	950	8	US-08-781-986A-1020	Sequence 1020, Ap	751	16	1.9	1748	13	US-10-117-722-712	Sequence 712, App
C 679	16	1.9	985	13	US-10-103-959A-4	Sequence 4, Appl1	752	16	1.9	1748	15	US-10-037-270-712	Sequence 712, App
C 680	16	1.9	1001	12	US-10-170-097-456	Sequence 456, App	753	16	1.9	1757	10	US-09-731-872-76	Sequence 76, Appl
C 681	16	1.9	1003	10	US-09-833-381-917	Sequence 917, App	754	16	1.9	1757	13	US-09-876-997-76	Sequence 76, Appl
C 682	16	1.9	1026	13	US-10-032-585-6491	Sequence 6491, Ap	755	16	1.9	1768	12	US-10-108-260A-2237	Sequence 2237, Ap
C 683	16	1.9	1038	13	US-10-087-887-15	Sequence 15, Appl	756	16	1.9	1791	15	US-10-198-846-10945	Sequence 10945, A
C 684	16	1.9	1059	12	US-09-938-842A-134	Sequence 134, App	757	16	1.9	1798	9	US-10-104-047-1537	Sequence 1537, Ap
C 685	16	1.9	1059	12	US-09-938-842A-134	Sequence 134, App	758	16	1.9	1815	12	US-09-925-302-157	Sequence 157, App
C 686	16	1.9	1094	13	US-09-814-353-19614	Sequence 19614, A	759	16	1.9	1833	10	US-09-938-842A-5346	Sequence 5346, Ap
C 687	16	1.9	1119	12	US-10-369-493-24745	Sequence 24745, A	760	16	1.9	1853	12	US-09-938-842A-5346	Sequence 5346, Ap
688	16	1.9	1132	13	US-10-027-632-261540	Sequence 261540, A	C 761	16	1.9	1855	13	US-10-027-632-212126	Sequence 212126, A
689	16	1.9	1132	13	US-10-027-632-261540	Sequence 261540, A	C 762	16	1.9	1855	14	US-10-027-632-212126	Sequence 212126, A
C 690	16	1.9	1134	14	US-09-974-300-128	Sequence 128, App	C 763	16	1.9	1859	10	US-09-915-582-23	Sequence 23, Appl
C 691	16	1.9	1139	14	US-10-027-632-119463	Sequence 119463, A	C 764	16	1.9	1859	12	US-10-277-802-23	Sequence 23, Appl
C 692	16	1.9	1139	14	US-10-027-632-119463	Sequence 119463, A	765	16	1.9	1859	12	US-10-369-493-25948	Sequence 25948, A
693	16	1.9	1159	10	US-09-925-300-257	Sequence 257, App	766	16	1.9	1880	12	US-10-108-260A-2064	Sequence 2064, Ap
694	16	1.9	1164	13	US-10-240-965-179	Sequence 179, App	767	16	1.9	1903	13	US-09-873-047-239	Sequence 239, App
695	16	1.9	1179	10	US-09-764-877-2814	Sequence 2814, Ap	C 768	16	1.9	1907	12	US-10-104-047-257	Sequence 257, App
696	16	1.9	1179	10	US-09-764-877-2815	Sequence 2814, Ap	C 769	16	1.9	1934	12	US-10-062-674-1514	Sequence 1514, Ap
697	16	1.9	1179	10	US-09-938-842A-44	Sequence 44, Appl	770	16	1.9	1930	10	US-09-731-872-114	Sequence 174, Ap
698	16	1.9	1179	12	US-09-938-842A-44	Sequence 44, Appl	771	16	1.9	1930	10	US-09-876-997-174	Sequence 174, Ap
699	16	1.9	1179	12	US-10-242-515-2814	Sequence 2814, Ap	C 772	16	1.9	1995	12	US-10-310-154-168	Sequence 168, App
700	16	1.9	1182	13	US-10-242-515-2815	Sequence 2815, Ap	C 773	16	1.9	2000	10	US-09-938-842A-2946	Sequence 2946, Ap
701	16	1.9	1182	13	US-10-259-165-155	Sequence 155, App	C 774	16	1.9	2000	10	US-09-938-842A-4834	Sequence 4834, Ap
C 702	16	1.9	1214	13	US-10-027-632-203565	Sequence 203565, A	C 775	16	1.9	2000	10	US-09-938-842A-4834	Sequence 4834, Ap
C 703	16	1.9	1214	14	US-10-027-632-203565	Sequence 203565, A	C 776	16	1.9	2000	12	US-09-938-842A-2946	Sequence 2946, Ap
C 704	16	1.9	1216	14	US-10-027-632-210428	Sequence 210428, A	C 777	16	1.9	2000	12	US-09-938-842A-4834	Sequence 4834, Ap
C 705	16	1.9	1216	14	US-10-027-632-210428	Sequence 210428, A	C 778	16	1.9	2000	12	US-10-260-238-2554	Sequence 2554, Ap
706	16	1.9	1220	9	US-09-962-436-301	Sequence 301, App	779	16	1.9	2000	12	US-10-260-238-2554	Sequence 2554, Ap
707	16	1.9	1222	13	US-10-027-632-250983	Sequence 250983, A	C 780	16	1.9	2001	15	US-10-172-086-62	Sequence 62, Appl
708	16	1.9	1222	14	US-10-027-632-250983	Sequence 250983, A	781	16	1.9	2014	15	US-10-099-395-4	Sequence 4, Appl1
709	16	1.9	1227	12	US-10-369-493-42125	Sequence 42125, A	C 782	16	1.9	2038	12	US-10-369-493-27284	Sequence 27284, A
710	16	1.9	1252	15	US-10-043-487-81	Sequence 81, Appl	C 783	16	1.9	2038	12	US-10-104-047-1920	Sequence 1920, Ap
711	16	1.9	1258	12	US-10-264-237-427	Sequence 427, App	C 784	16	1.9	2058	13	US-10-087-887-117	Sequence 17, Appl
C 712	16	1.9	1297	12	US-10-264-049-192	Sequence 192, App	785	16	1.9	2115	9	US-09-815-242-9867	Sequence 9867, Ap
713	16	1.9	1302	9	US-09-815-242-636	Sequence 436, Ap	786	16	1.9	2130	12	US-10-108-260A-1749	Sequence 1749, Ap
714	16	1.9	1302	9	US-09-815-242-6592	Sequence 8592, Ap	787	16	1.9	2202	12	US-10-369-493-26173	Sequence 26173, A
715	16	1.9	1307	14	US-10-027-632-204758	Sequence 204758, A	788	16	1.9	2202	10	US-10-369-493-46508	Sequence 46508, A
C 716	16	1.9	1307	14	US-10-027-632-204758	Sequence 204758, A	789	16	1.9	2220	12	US-09-974-300-102	Sequence 102, App
C 717	16	1.9	1308	14	US-10-062-254-151	Sequence 151, App	790	16	1.9	2227	13	US-10-094-749-518	Sequence 518, App
718	16	1.9	1332	10	US-09-938-842A-1753	Sequence 1753, Ap	791	16	1.9	2227	13	US-10-006-285-243	Sequence 243, App
719	16	1.9	1332	12	US-09-938-842A-1753	Sequence 1753, Ap	C 792	16	1.9	2228	12	US-10-104-047-1314	Sequence 1314, Ap
C 720	16	1.9	1355	15	US-10-050-882-16	Sequence 16, Appl	793	16	1.9	2233	10	US-09-938-842A-2160	Sequence 2160, Ap
C 721	16	1.9	1422	10	US-09-764-877-3596	Sequence 3596, Ap	794	16	1.9	2233	12	US-09-938-842A-2160	Sequence 2160, Ap
C 722	16	1.9	1422	12	US-10-242-515-3596	Sequence 3596, Ap	C 795	16	1.9	2270	11	US-09-764-891-7579	Sequence 7279, Ap
723	16	1.9	1434	12	US-10-369-493-44958	Sequence 44958, A	C 796	16	1.9	2270	11	US-09-764-891-7581	Sequence 7381, Ap
724	16	1.9	1466	13	US-09-814-353-19226	Sequence 19226, A	797	16	1.9	2334	13	US-10-252-197-310	Sequence 310, App
C 725	16	1.9	1476	13	US-09-738-269-16	Sequence 16, Appl	798	16	1.9	2339	12	US-10-369-493-25799	Sequence 25799, A
C 726	16	1.9	1476	14	US-10-023-437-16	Sequence 16, Appl	C 799	16	1.9	2336	13	US-10-027-632-263172	Sequence 263172, A
727	16	1.9	1503	12	US-10-369-493-41223	Sequence 41223, A	C 800	16	1.9	2336	14	US-10-027-632-263172	Sequence 263172, A
728	16	1.9	1504	12	US-10-369-493-30083	Sequence 30083, A	C 801	16	1.9	2337	15	US-10-183-116-80	Sequence 80, Appl
729	16	1.9	1508	15	US-10-106-698-528	Sequence 528, App	802	16	1.9	2427	10	US-09-954-531-573	Sequence 573, App
730	16	1.9	1518	13	US-10-168-097A-11	Sequence 11, Appl	803	16	1.9	2433	9	US-09-815-242-9538	Sequence 9538, Ap
C 731	16	1.9	1525	11	US-09-764-891-6547	Sequence 6547, Ap	804	16	1.9	2433	13	US-09-769-744A-73	Sequence 73, Appl
C 732	16	1.9	1525	15	US-10-205-428-744	Sequence 744, App	805	16	1.9	2437	14	US-10-096-264-1	Sequence 1, Appl1
C 733	16	1.9	1558	13	US-10-027-632-259994	Sequence 259994, A	806	16	1.9	2437	14	US-09-934-455-449	Sequence 449, App
734	16	1.9	1558	13	US-10-027-632-259995	Sequence 259995, A	C 807	16	1.9	2435	13	US-10-027-632-265068	Sequence 265068, A
735	16	1.9	1558	13	US-10-027-632-259995	Sequence 259995, A	C 808	16	1.9	2435	13	US-10-027-632-265069	Sequence 265069, A
736	16	1.9	1558	13	US-10-027-632-259997	Sequence 259997, A	C 809	16	1.9	2435	13	US-10-027-632-265070	Sequence 265070, A
737	16	1.9	1558	14	US-10-027-632-259994	Sequence 259994, A	C 810	16	1.9	2435	14	US-10-027-632-265068	Sequence 265068, A
738	16	1.9	1558	14	US-10-027-632-259994	Sequence 259994, A	C 811	16	1.9	2435	14	US-10-027-632-265069	Sequence 265069, A
739	16	1.9	1558	14	US-10-027-632-259996	Sequence 259996, A	C 812	16	1.9	2435	14	US-10-027-632-265070	Sequence 265070, A
740	16	1.9	1558	14	US-10-027-632-259997	Sequence 259997, A	813	16	1.9	2510	12	US-10-108-260A-1898	Sequence 1898, Ap
741	16	1.9	1607	14	US-10-096-241-3	Sequence 3, Appl1	814	16	1.9	2521	15	US-10-198-846-13275	Sequence 13275, A
742	16	1.9	1626	10	US-09-974-300-4490	Sequence 4490, Ap	815	16	1.9	2521	13	US-10-117-722-1302	Sequence 1302, App
743	16	1.9	1635	13	US-10-289-161A-4	Sequence 4, Appl1	816	16	1.9	2521	15	US-10-037-270-302	Sequence 302, App
744	16	1.9	1635	13	US-10-289-161A-5	Sequence 5, Appl1	817	16	1.9	2528	12	US-10-369-493-36618	Sequence 36618, A
745	16	1.9	1635	13	US-10-293-582-20	Sequence 20, Appl	818	16	1.9	2606	12	US-10-104-047-17	Sequence 17, Appl

819	16	1.9	2659	10	US-09-981-876-42	Sequence 42, App1	C 892	16	1.9	6954	13	US-10-311-455-1363	Sequence 1363, Ap
820	16	1.9	2659	11	US-09-148-545-42	Sequence 42, App1	893	16	1.9	7488	11	US-09-764-891-9489	Sequence 9489, Ap
C 821	16	1.9	2659	12	US-10-108-2608-616	Sequence 616, App	894	16	1.9	7488	11	US-09-764-891-9489	Sequence 10082, A
822	16	1.9	2681	10	US-09-917-8008-1458	Sequence 1458, Ap	895	16	1.9	7488	15	US-10-205-428-870	Sequence 870, App
823	16	1.9	2826	9	US-09-815-242-4558	Sequence 458, Ap	896	16	1.9	7488	15	US-10-205-428-992	Sequence 992, App
824	16	1.9	2847	9	US-09-815-242-8413	Sequence 8413, Ap	897	16	1.9	7488	15	US-10-205-428-992	Sequence 1757, Ap
825	16	1.9	2873	12	US-10-108-2608-1884	Sequence 1884, Ap	898	16	1.9	7566	10	US-09-728-952-1	Sequence 1, App1
826	16	1.9	2910	13	US-10-193-764-31	Sequence 31, App1	899	16	1.9	7633	13	US-10-311-455-2153	Sequence 2153, Ap
827	16	1.9	2917	13	US-09-814-353-19631	Sequence 19631, A	900	16	1.9	7633	13	US-09-728-952-1	Sequence 537, App
828	16	1.9	2928	13	US-10-193-764-29	Sequence 29, App1	901	16	1.9	7633	10	US-09-070-927A-328	Sequence 228, App
829	16	1.9	2938	15	US-10-198-846-9862	Sequence 9862, Ap	902	16	1.9	7893	10	US-10-252-157-411	Sequence 411, App
C 830	16	1.9	2981	13	US-10-096-534-45	Sequence 45, App1	903	16	1.9	8131	13	US-10-311-455-1955	Sequence 1955, Ap
831	16	1.9	3001	13	US-10-147-603-116	Sequence 116, App	904	16	1.9	8552	13	US-10-311-455-1955	Sequence 1, App1
832	16	1.9	3201	13	US-10-032-585-69312	Sequence 69312, Ap	C 905	16	1.9	10558	12	US-10-372-683-19	Sequence 19, App1
C 833	16	1.9	3234	13	US-10-032-585-6931	Sequence 6931, Ap	C 906	16	1.9	10558	12	US-10-096-534-14	Sequence 14, App1
C 834	16	1.9	3276	13	US-10-027-632-114370	Sequence 114370, A	C 907	16	1.9	10677	13	US-09-814-353-19948	Sequence 19948, A
C 835	16	1.9	3276	14	US-10-027-632-114370	Sequence 114370, A	C 908	16	1.9	10766	15	US-10-084-817-43	Sequence 43, App1
C 836	16	1.9	3355	12	US-10-159-563-351	Sequence 351, App	C 909	16	1.9	11115	9	US-09-908-711-166	Sequence 166, App
C 837	16	1.9	3381	9	US-09-815-242-4318	Sequence 4318, Ap	C 910	16	1.9	11260	13	US-10-240-453-28	Sequence 28, App1
C 838	16	1.9	3477	9	US-09-815-242-4424	Sequence 4424, Ap	C 911	16	1.9	11260	13	US-10-239-676-20	Sequence 20, App1
C 839	16	1.9	3496	10	US-09-974-300-2632	Sequence 2632, Ap	C 912	16	1.9	12541	10	US-09-764-877-2184	Sequence 2184, Ap
C 840	16	1.9	3613	13	US-10-027-632-261421	Sequence 261421, A	C 913	16	1.9	12541	12	US-10-242-515-2184	Sequence 2184, Ap
C 841	16	1.9	3613	13	US-10-027-632-261422	Sequence 261422, A	C 914	16	1.9	13086	8	US-08-781-9864-16	Sequence 16, App1
C 842	16	1.9	3613	14	US-10-027-632-261421	Sequence 261421, A	C 915	16	1.9	13125	13	US-10-311-455-1200	Sequence 1200, App
C 843	16	1.9	3800	12	US-10-369-493-27429	Sequence 27429, A	C 916	16	1.9	13125	13	US-10-240-485-110	Sequence 110, App
844	16	1.9	3858	12	US-10-006-285-429	Sequence 429, App	C 917	16	1.9	15734	9	US-09-764-869-1922	Sequence 1922, Ap
845	16	1.9	4031	12	US-10-264-049-952	Sequence 952, App	C 918	16	1.9	15734	12	US-10-227-577-1922	Sequence 1922, Ap
C 846	16	1.9	4137	13	US-10-341-434-7	Sequence 7, App1	C 919	16	1.9	15734	15	US-10-091-504-1822	Sequence 1822, Ap
847	16	1.9	4158	13	US-09-764-891-6570	Sequence 6570, Ap	C 920	16	1.9	17220	13	US-10-311-455-628	Sequence 628, App
848	16	1.9	4158	13	US-10-259-165-243	Sequence 243, App	C 921	16	1.9	17220	13	US-09-070-927A-83	Sequence 83, App1
849	16	1.9	4165	11	US-09-764-891-6571	Sequence 6571, Ap	C 922	16	1.9	17588	13	US-10-311-455-574	Sequence 574, App
850	16	1.9	4189	12	US-10-369-493-27741	Sequence 27741, A	C 923	16	1.9	17966	12	US-09-764-877-2695	Sequence 2695, Ap
851	16	1.9	4316	12	US-09-880-107-3713	Sequence 3713, A	C 924	16	1.9	17966	12	US-10-242-515-2695	Sequence 2695, Ap
C 852	16	1.9	4343	12	US-10-159-563-184	Sequence 184, App	C 925	16	1.9	18355	8	US-08-781-9864-67	Sequence 67, App1
C 853	16	1.9	4425	13	US-10-295-681-56	Sequence 56, App	C 926	16	1.9	18981	11	US-09-764-891-7649	Sequence 7649, Ap
C 854	16	1.9	4425	13	US-10-295-681-56	Sequence 56, App	C 927	16	1.9	19575	8	US-10-311-455-1423	Sequence 1423, Ap
C 855	16	1.9	4425	13	US-10-295-681-58	Sequence 58, App	C 928	16	1.9	29685	10	US-09-752-820A-3	Sequence 3, App1
C 856	16	1.9	4517	13	US-10-007-926A-304	Sequence 304, App	C 929	16	1.9	32188	9	US-09-813-318A-3	Sequence 3, App1
C 857	16	1.9	4517	16	US-10-298-377A-3	Sequence 3, App1	C 930	16	1.9	32188	9	US-09-764-860-799	Sequence 799, App
C 858	16	1.9	4543	14	US-10-098-841-173	Sequence 173, App	C 931	16	1.9	32188	15	US-10-212-87-799	Sequence 799, App
C 859	16	1.9	4808	8	US-08-781-986A-200	Sequence 200, App	C 932	16	1.9	32188	15	US-10-074-095-799	Sequence 799, App
C 860	16	1.9	5007	13	US-10-037-598-34	Sequence 34, App1	C 933	16	1.9	32480	13	US-09-847-101B-23	Sequence 23, App1
C 861	16	1.9	5077	14	US-10-047-412A-24	Sequence 24, App1	C 934	16	1.9	32788	13	US-09-482-682-27	Sequence 27, App1
C 862	16	1.9	5128	14	US-10-098-841-66	Sequence 66, App1	C 935	16	1.9	33622	12	US-10-351-890-44	Sequence 44, App1
C 863	16	1.9	5283	13	US-10-311-455-1866	Sequence 1866, Ap	C 936	16	1.9	33622	12	US-10-403-337-44	Sequence 44, App1
C 864	16	1.9	5344	14	US-10-098-841-65	Sequence 65, App1	C 937	16	1.9	33622	12	US-10-383-846-5	Sequence 5, App1
C 865	16	1.9	5443	13	US-10-311-455-856	Sequence 856, App	C 938	16	1.9	34053	15	US-10-114-170-96	Sequence 96, App1
C 866	16	1.9	5519	15	US-10-354-958-69	Sequence 11, App1	C 939	16	1.9	34457	11	US-09-111-911-5	Sequence 5, App1
C 867	16	1.9	5547	15	US-10-285-649A-7	Sequence 16, App1	C 940	16	1.9	34555	12	US-10-313-986-479	Sequence 479, App
C 868	16	1.9	5547	15	US-10-285-649A-7	Sequence 16, App1	C 941	16	1.9	34555	12	US-10-117-982-479	Sequence 479, App
C 869	16	1.9	5547	15	US-10-285-649A-7	Sequence 16, App1	C 942	16	1.9	35211	12	US-10-351-890-43	Sequence 43, App1
C 870	16	1.9	5547	15	US-10-285-649A-7	Sequence 16, App1	C 943	16	1.9	35211	12	US-10-351-890-43	Sequence 43, App1
C 871	16	1.9	5582	13	US-09-917-800A-1459	Sequence 1459, Ap	C 944	16	1.9	35211	12	US-10-403-337-43	Sequence 43, App1
C 872	16	1.9	5675	13	US-10-311-455-942	Sequence 942, App	C 945	16	1.9	35408	15	US-10-155-640-3	Sequence 3, App1
C 873	16	1.9	5675	13	US-10-311-455-942	Sequence 942, App	C 946	16	1.9	35871	9	US-09-956-335-2	Sequence 2, App1
C 874	16	1.9	5789	12	US-10-402-365-1	Sequence 1, App1	C 947	16	1.9	35935	10	US-09-725-720-43	Sequence 43, App1
C 875	16	1.9	6047	13	US-10-311-455-1431	Sequence 1431, Ap	C 948	16	1.9	35935	10	US-09-782-378A-4	Sequence 4, App1
C 876	16	1.9	6107	10	US-09-870-759-102	Sequence 102, App	C 949	16	1.9	35935	10	US-09-782-378A-4	Sequence 4, App1
C 877	16	1.9	6107	10	US-09-870-759-102	Sequence 102, App	C 950	16	1.9	35935	10	US-09-782-378A-4	Sequence 4, App1
C 878	16	1.9	6112	12	US-10-242-515-2958	Sequence 2958, Ap	C 951	16	1.9	35935	10	US-09-739-007-43	Sequence 43, App1
C 879	16	1.9	6156	13	US-10-204-708-59	Sequence 59, App1	C 952	16	1.9	35937	10	US-09-782-378A-3	Sequence 3, App1
C 880	16	1.9	6271	11	US-09-764-891-7626	Sequence 7626, Ap	C 953	16	1.9	35937	10	US-09-782-378A-3	Sequence 3, App1
C 881	16	1.9	6277	13	US-10-311-455-2013	Sequence 2013, Ap	C 954	16	1.9	43442	12	US-10-085-117-7	Sequence 7, App1
C 882	16	1.9	6609	13	US-10-311-455-1856	Sequence 1856, Ap	C 955	16	1.9	43442	12	US-10-085-117-7	Sequence 7, App1
C 883	16	1.9	6766	15	US-09-764-847-1878	Sequence 1878, Ap	C 956	16	1.9	49984	9	US-09-739-457-5	Sequence 5, App1
C 884	16	1.9	6766	15	US-10-092-154-1878	Sequence 1878, Ap	C 957	16	1.9	66109	10	US-09-880-107-3768	Sequence 3768, Ap
C 885	16	1.9	6952	13	US-10-133-013-144	Sequence 144, App	C 958	16	1.9	72480	12	US-09-873-367C-1060	Sequence 2, App1
C 886	16	1.9	6952	13	US-10-133-013-144	Sequence 144, App	C 959	16	1.9	80246	11	US-09-728-552-4	Sequence 4, App1
C 887	16	1.9	6952	13	US-10-133-013-144	Sequence 144, App	C 960	16	1.9	80246	11	US-09-728-552-4	Sequence 4, App1
C 888	16	1.9	6952	13	US-10-133-013-144	Sequence 144, App	C 961	16	1.9	80246	11	US-09-728-552-4	Sequence 4, App1
C 889	16	1.9	6952	13	US-10-133-013-144	Sequence 144, App	C 962	16	1.9	80246	11	US-09-728-552-4	Sequence 4, App1
C 890	16	1.9	6952	13	US-10-133-013-144	Sequence 144, App	C 963	16	1.9	80246	11	US-09-728-552-4	Sequence 4, App1
C 891	16	1.9	6952	13	US-10-133-013-144	Sequence 144, App	C 964	16	1.9	80246	11	US-09-728-552-4	Sequence 4, App1

```
665 16 1.9 80595 11 US-09-728-552-3 Sequence 3, Appl1
666 16 1.9 90541 9 US-09-759-359A-3 Sequence 3, Appl1
667 16 1.9 90541 13 US-10-207-973-3 Sequence 3, Appl1
668 16 1.9 99998 12 US-10-085-117-34 Sequence 34, Appl1
669 16 1.9 104000 15 US-10-012-984-14 Sequence 11, Appl1
670 16 1.9 118518 12 US-10-189-256-11 Sequence 11, Appl1
671 16 1.9 118518 12 US-10-189-256-11 Sequence 11, Appl1
672 16 1.9 174566 15 US-10-020-141-1 Sequence 6, Appl1
673 16 1.9 180216 9 US-09-835-232-6 Sequence 6, Appl1
674 16 1.9 180216 13 US-10-308-485-6 Sequence 142, App
675 16 1.9 193357 12 US-10-085-117-142 Sequence 3814, Ap
676 16 1.9 198285 10 US-09-880-107-3814 Sequence 286, App
677 16 1.9 202802 12 US-10-085-117-286 Sequence 3, Appl1
678 16 1.9 254366 11 US-09-822-871-3 Sequence 23, Appl1
679 16 1.9 260209 13 US-10-265-071-23 Sequence 23, Appl1
680 16 1.9 260209 13 US-10-265-071-23 Sequence 23, Appl1
681 16 1.9 335913 11 US-09-754-853A-2 Sequence 2, Appl1
682 16 1.9 335913 11 US-09-754-853A-2 Sequence 3, Appl1
683 16 1.9 402850 11 US-09-844-653-5 Sequence 5, Appl1
684 16 1.9 580073 13 US-10-205-220-1 Sequence 1, Appl1
685 16 1.9 659158 13 US-09-771-208-20 Sequence 20, Appl1
686 16 1.9 713059 13 US-10-027-632-174581 Sequence 174581,
687 16 1.9 713059 14 US-10-027-632-174581 Sequence 174581,
688 16 1.9 744802 12 US-10-292-798-1369 Sequence 1369, Ap
689 16 1.9 786431 13 US-10-412-277-3 Sequence 3, Appl1
690 16 1.9 832900 12 US-10-292-798-1393 Sequence 1393, Ap
691 16 1.9 1503841 9 US-09-795-668-1 Sequence 1, Appl1
692 16 1.9 1503841 10 US-09-946-807-1 Sequence 1, Appl1
693 16 1.9 1503841 13 US-10-312-841-1 Sequence 1, Appl1
694 16 1.9 3673778 13 US-10-312-841-2 Sequence 2, Appl1
695 16 1.9 3673778 13 US-10-045-360-18 Sequence 18, Appl1
696 15 1.8 25 13 US-10-098-2638-9693 Sequence 9693, Ap
697 15 1.8 25 15 US-10-098-2638-9693 Sequence 91753, A
698 15 1.8 25 15 US-10-098-2638-91753 Sequence 91753, A
699 15 1.8 25 15 US-10-098-2638-91989 Sequence 355, App
C1000 15 1.8 51 10 US-09-922-261-355
```

## ALIGNMENTS

```
RESULT 1
US-10-033-190-1
; Sequence 1, Application US/10033190
; Publication No. US20020133848A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis Plant Sciences, Inc.
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF AN ANTHOCYANIN MUTANT (ANT)
; FILE REFERENCE: EP01-002C
; CURRENT APPLICATION NUMBER: US/10/033,190
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,685
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-10-033-190-1
```

```
Query Match 100.0%; Score 825; DB 14; Length 825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ATGAACAGTACATCTATCTTTCATTTGGAGTGAAGAAAGTTCATGAGTGAAGAA 60
Db 1 ATGAACAGTACATCTATCTTTCATTTGGAGTGAAGAAAGTTCATGAGTGAAGAA 60
Qy 61 GATTTCTTCTAGAGAAATGATTAAGTATGATGAGTGAAGAAATGAGCATCTGTTCC 120
Db 61 GATTTCTTCTAGAGAAATGATTAAGTATGATGAGTGAAGAAATGAGCATCTGTTCC 120
```

```
Qy 121 ATAGAGCTGCTCTGAATAGATGTCGGAAGAGTTGATGAGTGGCTGAATATCTA 180
Db 121 ATAGAGCTGCTCTGAATAGATGTCGGAAGAGTTGATGAGTGGCTGAATATCTA 180
Qy 181 AGGCCATATCAAGAGAGTGACTTTGAACAAGATGAATGATTTCAATTTTGGAGCTT 240
Db 181 AGGCCATATCAAGAGAGTGACTTTGAACAAGATGAATGATTTCAATTTTGGAGCTT 240
Qy 241 CATTAAGCTTTAGGCAACAGATGCTATTTGCTGATGATCTCCGGAAGAGAGCT 300
Db 241 CATTAAGCTTTAGGCAACAGATGCTATTTGCTGATGATCTCCGGAAGAGAGCT 300
Qy 301 AAGCATGTGAAAACATTTGGAACACATTAATCTTTAAGAGAGTTAAATCTACTAAATTT 360
Db 301 AAGCATGTGAAAACATTTGGAACACATTAATCTTTAAGAGAGTTAAATCTACTAAATTT 360
Qy 361 GTTCTCTGCGAAGAAATTAACAATAAGTGTGAGAAATTAATGATTAAGTGAATTAAT 420
Db 361 GTTCTCTGCGAAGAAATTAACAATAAGTGTGAGAAATTAATGATTAAGTGAATTAAT 420
Qy 421 AAACCTCAAGCAAGATATTTCTAGACACATGAGAAATGTAACAATTAATGTA 480
Db 421 AAACCTCAAGCAAGATATTTCTAGACACATGAGAAATGTAACAATTAATGTA 480
Qy 481 ATTTGAGCAGAGAGAAACATTTGCAAGAAATTAATGATGAGAAACAACTCCAGATGCA 540
Db 481 ATTTGAGCAGAGAGAAACATTTGCAAGAAATTAATGATGAGAAACAACTCCAGATGCA 540
Qy 541 TCGATGACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 541 TCGATGACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Qy 601 GAGGAGATGAAGAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
Db 601 GAGGAGATGAAGAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
Qy 661 GAAATATCACACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
Db 661 GAAATATCACACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
Qy 721 CATGAATAATGGGATGATTTCTTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
Db 721 CATGAATAATGGGATGATTTCTTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
Qy 781 GATGATTTTCTGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 825
Db 781 GATGATTTTCTGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 825
```

```
RESULT 2
US-10-033-190-4
; Sequence 4, Application US/10033190
; Publication No. US20020133848A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis Plant Sciences, Inc.
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF AN ANTHOCYANIN MUTANT (ANT)
; FILE REFERENCE: EP01-002C
; CURRENT APPLICATION NUMBER: US/10/033,190
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,685
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1012
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-10-033-190-4
```

```
Query Match 68.7%; Score 567; DB 14; Length 1012;
Best Local Similarity 100.0%; Pred. No. 1,4e-277;
```



Matches	567: Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
Qy	259	AGATGCTACTATTAATGCTGCTGTAAGACTTCCCGGAGAGACAGCTTAAGATGTGAAAAACTAT	318					
Db	446	AGATGCTACTATTAATGCTGCTGTAAGACTTCCCGGAGAGACAGCTTAAGATGTGAAAAACTAT	505					
Qy	319	TGGAACACTAATCTCTTAAGAAAGTTAAATACTACTAAAAATTGTCCTCGCGAAAAGATT	378					
Db	506	TGGAACACTAATCTCTTAAGAAAGTTAAATACTACTAAAAATTGTCCTCGCGAAAAGATT	565					
Qy	379	AACAATAAGTGTGAGAAATTAGTACTAAGATTGAAATTATATAAACTTCACAGACGCAAG	438					
Db	566	AACAATAAGTGTGAGAAATTAGTACTAAGATTGAAATTATATAAACTTCACAGACGCAAG	625					
Qy	439	TATTTCTCAAGCACAATGAGAAATGTACAAACAATATGTATTTTGGACGAGAGGAA	498					
Db	626	TATTTCTCAAGCACAATGAGAAATGTACAAACAATATGTATTTTGGACGAGAGGAA	685					
Qy	499	CATTGCAAGGAAATTAATAGTAGAGAAACAACTCCAGATGCATGCATGGACAAACGTAGAT	558					
Db	686	CATTGCAAGGAAATTAATAGTAGAGAAACAACTCCAGATGCATGCATGGACAAACGTAGAT	745					
Qy	559	CCATGCTGATTAATTTTACTGGAATAATGCAATGACGATTTGAAGAAGATGAAGAGTT	618					
Db	746	CCATGCTGATTAATTTTACTGGAATAATGCAATGACGATTTGAAGAAGATGAAGAGTT	805					
Qy	619	GTAATTAATTAATGAAAAAAACACTTAACAAGTTTGTATCATGAGAAATATCACCACTTA	678					
Db	806	GTAATTAATTAATGAAAAAAACACTTAACAAGTTTGTATCATGAGAAATATCACCACTTA	865					
Qy	679	AATATGTGTGAAGTAACTCTCCATGCAACAAGACAAATTAAGTCATGAAAAATTGGGGTGA	738					
Db	866	AATATGTGTGAAGTAACTCTCCATGCAACAAGACAAATTAAGTCATGAAAAATTGGGGTGA	925					
Qy	739	TTTTCTCTTAATTTTACCAACCCATGCAACAAGAGTACAAAATGATGATTTTCTGCTGA	798					
Db	926	TTTTCTCTTAATTTTACCAACCCATGCAACAAGAGTACAAAATGATGATTTTCTGCTGA	985					
Qy	799	ATTGACTTATGGAATCTACTTGATTAA	825					
Db	986	ATTGACTTATGGAATCTACTTGATTAA	1012					
RESULT 3								
US-10-021-811-45								
; Sequence 45, Application US/10021811								
; Publication No. US20030024007A1								
GENERAL INFORMATION:								
; APPLICANT: Cahoon, Rebecca E.								
; APPLICANT: Fang, Yiwen								
; APPLICANT: Odell, Joan								
; APPLICANT: Meng, Zude								
TITLE OF INVENTION: Plant Myb Transcription Factor Homologs								
FILE REFERENCE: BBI294 US NA								
CURRENT APPLICATION NUMBER: US/10/021,811								
CURRENT FILING DATE: 2001-12-14								
PRIOR APPLICATION NUMBER: 60/110,609								
PRIOR FILING DATE: 1998-December-02								
NUMBER OF SEQ ID NOS: 63								
SOFTWARE: Microsoft Office 97								
SEQ ID NO 45								
LENGTH: 1236								
TYPE: DNA								
ORGANISM: Glycine max								
FEATURE:								
NAME/KEY: unsure								
LOCATION: (519)								
NAME/KEY: unsure								
LOCATION:								

```

NAME/KEY: unsure
LOCATION: (800)
NAME/KEY: unsure
LOCATION: (1124)
NAME/KEY: unsure
LOCATION: (1151)
US-10-021-811-45

Query Match
Best Local Similarity 100.0%; Pred. NO. 0.52;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2.8%; Score 23; DB 15; Length 1236;

157 AGATTGAGTGGCTGAATTATCT 179
|||||
282 AGATTGAGTGGCTGAATTATCT 304

RESULT 4
US-09-934-455-251
; Sequence 251, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubbell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omalra
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 251
; LENGTH: 1300
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (168)..(1196)
; OTHER INFORMATION: GI640
US-09-934-455-251

Query Match
Best Local Similarity 100.0%; Score 23; DB 11; Length 1300;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2.8%; Score 23; DB 11; Length 1300;

253 GGCACAGATGCTCACTTATTC 275
|||||
423 GGCACAGATGCTCACTTATTC 445

RESULT 5
US-10-021-811-43
; Sequence 43, Application US/10021811
; Publication No. US20030024007A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yifan
; APPLICANT: Odell, Joan

```

APPLICANT: Meng, Zude  
TITLE OF INVENTION: Plant Myb Transcription Factor Homologs  
FILE REFERENCE: B81294 US NA  
CURRENT APPLICATION NUMBER: US/10/021,811  
CURRENT FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: 60/110,609  
PRIOR FILING DATE: 1998-December-02  
NUMBER OF SEQ ID NOS: 63  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 43  
LENGTH: 1348  
TYPE: DNA  
ORGANISM: Glycine max  
US-10-021-811-43

Query Match 2.8%; Score 23; DB 15; Length 1348;  
Best Local Similarity 100.0%; Pred. No. 0.52; Indels 0; Gaps 0;  
Matches 23; Conservative 0; Mismatches 0;

Oy 157 AGATTGAGTGGCTGATTTATCT 179  
Db 266 AGATTGAGTGGCTGATTTATCT 288

RESULT 6  
US-09-878-574-2454  
Sequence 2454, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Thompson, Michael D.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(13401)B  
CURRENT APPLICATION NUMBER: US/09/878,574  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 09/333,535  
PRIOR FILING DATE: 1999-06-14  
NUMBER OF SEQ ID NOS: 15775  
SEQ ID NO 2454  
LENGTH: 390  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(390)  
OTHER INFORMATION: unsure at all n locations  
OTHER INFORMATION: Clone ID: LIB3028-024-Q1-B1-G6  
US-09-878-574-2454

Query Match 2.5%; Score 21; DB 10; Length 390;  
Best Local Similarity 100.0%; Pred. No. 4.9; Indels 0; Gaps 0;  
Matches 21; Conservative 0; Mismatches 0;

Oy 304 GATGTGAAAACCTATTGGAAC 324  
Db 300 GATGTGAAAACCTATTGGAAC 320

RESULT 7  
US-10-329-960-1/c  
Sequence 1, Application US/10329960  
Publication No. US2003009277A1  
GENERAL INFORMATION:  
APPLICANT: Fleischmann et al.  
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag  
FILE REFERENCE: p8186p1  
CURRENT APPLICATION NUMBER: US/10/329,960  
CURRENT FILING DATE: 2003-01-02  
PRIOR APPLICATION NUMBER: US 09/643,990  
PRIOR FILING DATE: 2000-08-23

PRIOR APPLICATION NUMBER: US 08/487,429  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: US 08/426,787  
PRIOR FILING DATE: 1995-04-21  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: Patent version 3.1  
SEQ ID NO 1  
LENGTH: 1830121  
TYPE: DNA  
ORGANISM: Haemophilus influenzae  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4747)..(4747)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (9921)..(9921)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (10150)..(10150)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (29298)..(29298)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36543)..(36543)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36551)..(36551)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36636)..(36636)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (40808)..(40810)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44416)..(44416)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44905)..(44905)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44975)..(44975)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (45593)..(45593)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (45732)..(45732)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (47036)..(47036)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (51334)..(51334)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (51602)..(51602)

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139310)..(139310)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)

Query Match      2.5%; Score 21; DB 15; Length 1630121;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      788 TTTCCTGTAATTGACTTAT 808
      |||||
Db      1174328 TTTCCTGTAATTGACTTAT 1174308

RESULT 8
US-09-938-842A-121
; Sequence 121, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22

```

```

; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 121
; LENGTH: 741
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-121

Query Match          2.4%; Score 20; DB 10; Length 741;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      247 CTCTTAGGCAACAGATGCTC 266
Db      250 CTCTTAGGCAACAGATGCTC 269

RESULT 9
US-09-938-842A-121
; Sequence 121, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/2227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 121
; LENGTH: 741
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-121

Query Match          2.4%; Score 20; DB 12; Length 741;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      247 CTCTTAGGCAACAGATGCTC 266
Db      250 CTCTTAGGCAACAGATGCTC 269

RESULT 10
US-10-369-493-24702
; Sequence 24702, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24702
; LENGTH: 1782
; TYPE: DNA
```

```

; ORGANISM: Archaeoglobus fulgidus
US-10-369-493-24702

Query Match          2.4%; Score 20; DB 12; Length 1782;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      598 ATTGAAGAGATGAGAGGT 617
Db      1761 ATTGAAGAGATGAGAGGT 1780

RESULT 11
US-10-098-263B-91990/c
; Sequence 91990, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Miltman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1.1
; SEQ ID NO 91990
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-91990

Query Match          2.3%; Score 19; DB 15; Length 25;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      203 ACTTTGAACAAGATGAAGT 221
Db      21  ACTTTGAACAAGATGAAGT 3

RESULT 12
US-10-027-632-31145
; Sequence 31145, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 31145
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-31145
```

Query Match 2.3%; Score 19; DB 13; Length 966;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 726 AAATGGGGTGAATTTCT 744  
|||||  
Db 26 AAATGGGGTGAATTTCT 44

RESULT 13  
US-10-027-632-31145  
; Sequence 31145, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 31145  
; LENGTH: 966  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-31145

Query Match 2.3%; Score 19; DB 14; Length 966;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 726 AAATGGGGTGAATTTCT 744  
|||||  
Db 26 AAATGGGGTGAATTTCT 44

RESULT 14  
US-10-027-632-101162/c  
; Sequence 101162, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 101162  
; LENGTH: 1106  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-101162

Query Match 2.3%; Score 19; DB 13; Length 1106;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 524 AACAACTCCAGATGCATC 542  
|||||  
Db 324 AACAACTCCAGATGCATC 306

RESULT 15  
US-10-027-632-101162/c  
; Sequence 101162, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 101162  
; LENGTH: 1106  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-101162

Query Match 2.3%; Score 19; DB 14; Length 1106;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 524 AACAACTCCAGATGCATC 542  
|||||  
Db 324 AACAACTCCAGATGCATC 306

Search completed: January 30, 2004, 11:45:43  
Job time : 327 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 30, 2004, 09:26:18 ; Search time 60 Seconds  
(without alignments)  
6069.018 Million cell updates/sec

Title: US-10-033-190-1

Perfect score: 825  
Sequence: 1 atgaacagctacatcatgtc.....tatgaatcctactgttaa 825

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	21	2.5 1830121	4 US-09-557-884-1	Sequence 1, Appl
C 2	21	2.5 1830121	4 US-09-643-990A-1	Sequence 1, Appl
C 3	19	2.3 1394	4 US-09-453-7028-30	Sequence 30, Appl
C 4	19	2.3 1911	4 US-09-212-247C-1	Sequence 1, Appl
C 5	18	2.2 999	4 US-09-699-266A-5	Sequence 5, Appl
C 6	18	2.2 1033	4 US-09-610-185C-3	Sequence 3, Appl
C 7	18	2.2 3767	3 US-08-928-941D-28	Sequence 28, Appl
C 8	18	2.2 3767	3 US-08-928-941D-30	Sequence 30, Appl
C 9	18	2.2 3767	4 US-09-280-590A-28	Sequence 28, Appl
C 10	18	2.2 3767	4 US-09-280-590A-30	Sequence 30, Appl
C 11	18	2.2 7515	4 US-09-328-352-1695	Sequence 1695, Ap
C 12	18	2.2 10917	3 US-08-926-842B-11	Sequence 11, Appl
C 13	17	2.1 271	4 US-09-313-294A-3585	Sequence 3585, Ap
C 14	17	2.1 750	4 US-09-328-352-1167	Sequence 1167, Ap
C 15	17	2.1 1476	3 US-08-753-007A-7	Sequence 7, Appl
C 16	17	2.1 1476	3 US-09-398-496-7	Sequence 7, Appl
C 17	17	2.1 1884	3 US-08-753-007A-5	Sequence 5, Appl
C 18	17	2.1 1884	3 US-09-398-496-5	Sequence 5, Appl
C 19	17	2.1 2268	3 US-08-753-007A-31	Sequence 31, Appl
C 20	17	2.1 2268	3 US-09-398-496-31	Sequence 31, Appl
C 21	17	2.1 2718	4 US-09-134-001C-945	Sequence 945, App
C 22	17	2.1 2834	3 US-09-305-384-6	Sequence 6, Appl
C 23	17	2.1 2859	5 PCT-US96-05320A-637	Sequence 637, App
C 24	17	2.1 2883	2 US-08-533-306A-7	Sequence 7, Appl
C 25	17	2.1 2883	2 US-08-742-923A-7	Sequence 7, Appl
C 26	17	2.1 2933	3 US-09-262-773-409	Sequence 209, Appl
C 27	17	2.1 3178	3 US-08-968-563-7	Sequence 7, Appl

28	17	2.1 3178	3 US-08-969-683A-7	Sequence 7, Appl
29	17	2.1 3178	4 US-09-297-928-3	Sequence 3, Appl
30	17	2.1 3252	4 US-08-476-102A-4	Sequence 4, Appl
31	17	2.1 3252	5 PCT-US96-05320A-714	Sequence 714, App
32	17	2.1 6235	3 US-09-305-384-5	Sequence 5, Appl
33	17	2.1 6235	3 US-09-525-160A-6	Sequence 6, Appl
34	17	2.1 6265	4 US-09-129-112-3	Sequence 3, Appl
35	17	2.1 6679	3 US-09-305-384-1	Sequence 1, Appl
36	17	2.1 6679	3 US-09-525-160B-5	Sequence 5, Appl
37	17	2.1 23071	3 US-09-262-773-210	Sequence 210, App
38	17	2.1 1664976	4 US-08-916-421B-1	Sequence 1, Appl
39	17	2.1 1664976	4 US-08-916-421B-1	Sequence 1, Appl
40	16	1.9 48	4 US-09-302-620B-68	Sequence 68, Appl
41	16	1.9 80	4 US-09-478-189-3	Sequence 3, Appl
42	16	1.9 255	4 US-09-480-921B-26	Sequence 26, Appl
43	16	1.9 266	4 US-09-313-294A-2340	Sequence 2340, Ap
44	16	1.9 292	4 US-09-702-705-1781	Sequence 1781, Ap
45	16	1.9 292	4 US-09-736-457-1781	Sequence 1781, Ap
46	16	1.9 298	4 US-08-651-155B-70	Sequence 70, Appl
47	16	1.9 298	4 US-09-194-036B-70	Sequence 70, Appl
48	16	1.9 312	4 US-09-313-294A-7124	Sequence 7124, Ap
49	16	1.9 347	4 US-09-495-050A-188	Sequence 188, App
50	16	1.9 429	4 US-09-134-001C-636	Sequence 636, App
51	16	1.9 618	3 US-09-385-982-218	Sequence 218, App
52	16	1.9 815	3 US-08-906-769-128	Sequence 128, App
53	16	1.9 815	3 US-08-906-616-128	Sequence 128, App
54	16	1.9 815	3 US-08-639-075A-128	Sequence 128, App
55	16	1.9 815	3 US-09-012-431-128	Sequence 128, App
56	16	1.9 815	3 US-09-012-692-128	Sequence 128, App
57	16	1.9 815	3 US-08-906-613-128	Sequence 128, App
58	16	1.9 843	2 US-08-458-657-1	Sequence 1, Appl
59	16	1.9 843	2 US-08-860-995-1	Sequence 1, Appl
60	16	1.9 843	5 PCT-US95-00532A-1	Sequence 5, Appl
61	16	1.9 933	4 US-09-107-532A-558	Sequence 558, App
62	16	1.9 939	4 US-09-328-352-313	Sequence 313, App
63	16	1.9 956	4 US-09-610-185C-1	Sequence 1, Appl
64	16	1.9 957	4 US-08-961-527-371	Sequence 371, App
65	16	1.9 958	1 US-08-879-561-2	Sequence 2, Appl
66	16	1.9 983	1 US-08-525-864A-3	Sequence 3, Appl
67	16	1.9 1001	4 US-09-641-638-456	Sequence 456, App
68	16	1.9 1077	4 US-09-328-352-333	Sequence 333, App
69	16	1.9 1122	4 US-09-328-352-1627	Sequence 1627, Ap
70	16	1.9 1143	4 US-09-673-397-1	Sequence 1, Appl
71	16	1.9 1143	4 US-09-673-397-5	Sequence 5, Appl
72	16	1.9 1143	4 US-09-673-397-7	Sequence 7, Appl
73	16	1.9 1173	4 US-09-328-352-787	Sequence 787, App
74	16	1.9 1275	4 US-09-134-001C-2113	Sequence 2113, App
75	16	1.9 1494	4 US-09-107-532A-2521	Sequence 2521, App
76	16	1.9 1497	2 US-08-860-577-11	Sequence 11, Appl
77	16	1.9 1587	4 US-09-328-352-1606	Sequence 1606, Ap
78	16	1.9 1607	3 US-08-753-007A-3	Sequence 3, Appl
79	16	1.9 1607	3 US-09-398-496-3	Sequence 3, Appl
80	16	1.9 1659	1 US-08-231-729B-3	Sequence 3, Appl
81	16	1.9 1659	1 US-08-231-729B-4	Sequence 4, Appl
82	16	1.9 1748	4 US-09-620-312D-712	Sequence 712, App
83	16	1.9 1773	3 US-08-943-731-215	Sequence 215, App
84	16	1.9 1966	4 US-09-328-352-2622	Sequence 2622, App
85	16	1.9 2053	1 US-08-231-729B-5	Sequence 5, Appl
86	16	1.9 2187	4 US-09-107-532A-1429	Sequence 1429, Ap
87	16	1.9 2317	3 US-09-134-001C-1429	Sequence 5, Appl
88	16	1.9 2457	3 US-08-753-007A-1	Sequence 1, Appl
89	16	1.9 2457	3 US-09-398-496-1	Sequence 1, Appl
90	16	1.9 2571	4 US-09-620-312D-302	Sequence 302, App
91	16	1.9 2552	4 US-09-328-352-2604	Sequence 2604, App
92	16	1.9 2703	4 US-09-328-352-3946	Sequence 3946, App
93	16	1.9 2910	4 US-09-206-942-31	Sequence 31, Appl
94	16	1.9 2928	4 US-09-206-942-29	Sequence 29, Appl
95	16	1.9 3001	4 US-09-539-333D-116	Sequence 116, App
96	16	1.9 3105	4 US-09-390-234-15	Sequence 15, Appl
97	16	1.9 3441	2 US-08-525-864A-1	Sequence 1, Appl
98	16	1.9 3974	2 US-08-467-504-3	Sequence 3, Appl
99	16	1.9 4515	4 US-09-328-352-2920	Sequence 2920, App
100	16	1.9 4517	5 PCT-US93-06251-83	Sequence 83, Appl

101	1.9	5077	4	US-09-480-921B-24	Sequence 24, Appl	174	15	1.8	402	4	US-09-107-532A-2337	Sequence 2337, Ap
C 102	1.9	6216	3	US-09-213-053-1	Sequence 1, Appl	175	15	1.8	426	3	US-08-908-643C-52	Sequence 52, Appl
C 103	1.9	6234	3	US-08-971-188-1	Sequence 1, Appl	176	15	1.8	453	4	US-09-134-001C-881	Sequence 881, App
C 104	1.9	6246	3	US-08-943-731-640	Sequence 640, App	177	15	1.8	466	3	US-09-489-868A-10	Sequence 10, Appl
C 105	1.9	6270	1	US-08-418-893D-25	Sequence 25, Appl	178	15	1.8	483	4	US-09-107-532A-2123	Sequence 2123, Ap
C 106	1.9	6790	1	US-08-418-893D-22	Sequence 22, Appl	179	15	1.8	498	3	US-08-908-643C-50	Sequence 50, Appl
107	1.9	8697	4	US-08-961-527-123	Sequence 123, App	180	15	1.8	507	4	US-09-737-300-1	Sequence 1, Appl
108	1.9	9046	1	US-08-227-536-1	Sequence 1, Appl	181	15	1.8	510	4	US-09-615-192A-112	Sequence 132, App
109	1.9	9046	5	PCT-US95-04682-1	Sequence 1, Appl	182	15	1.8	523	4	US-09-679-409-59	Sequence 59, Appl
110	1.9	9636	1	US-08-323-170B-1	Sequence 1, Appl	183	15	1.8	524	2	US-08-975-316-64	Sequence 64, Appl
111	1.9	9636	4	US-08-954-441-1	Sequence 1, Appl	184	15	1.8	524	4	US-09-615-192A-64	Sequence 64, Appl
C 112	1.9	10207	1	US-08-920-812-2	Sequence 2, Appl	185	15	1.8	540	1	US-08-117-083-18	Sequence 18, Appl
C 113	1.9	10207	1	US-08-920-827-2	Sequence 2, Appl	186	15	1.8	540	1	US-09-328-352-2417	Sequence 2417, Ap
C 114	1.9	10207	1	US-08-921-177-2	Sequence 2, Appl	187	15	1.8	552	3	US-08-617-860B-27	Sequence 27, Appl
C 115	1.9	10207	1	US-08-362-577C-2	Sequence 2, Appl	188	15	1.8	558	4	US-09-134-001C-456	Sequence 456, App
C 116	1.9	10207	2	US-08-920-828-2	Sequence 2, Appl	189	15	1.8	605	4	US-09-690-364-19	Sequence 19, Appl
C 117	1.9	32207	2	US-08-770-379-20	Sequence 20, Appl	190	15	1.8	624	4	US-09-134-001C-1651	Sequence 1651, Ap
C 118	1.9	32207	3	US-08-757-669A-20	Sequence 20, Appl	191	15	1.8	655	4	US-09-643-597-208	Sequence 208, App
C 119	1.9	32207	4	US-09-230-371A-20	Sequence 20, Appl	192	15	1.8	655	4	US-09-480-884A-208	Sequence 208, App
C 120	1.9	32798	4	US-09-604-694B-1	Sequence 1, Appl	193	15	1.8	655	4	US-09-542-615A-208	Sequence 208, App
C 121	1.9	34063	4	US-09-453-702B-96	Sequence 96, Appl	194	15	1.8	741	4	US-09-606-421B-208	Sequence 208, App
C 122	1.9	34303	2	US-08-735-609-4	Sequence 4, Appl	195	15	1.8	751	4	US-09-328-352-565	Sequence 565, App
C 123	1.9	34303	2	US-08-735-609-4	Sequence 4, Appl	196	15	1.8	751	4	US-09-280-116-86	Sequence 86, Appl
C 124	1.9	34303	3	US-09-315-372-4	Sequence 4, Appl	197	15	1.8	785	6	US-08-436-748-4	Sequence 86, Appl
C 125	1.9	34303	3	US-09-244-752-4	Sequence 4, Appl	198	15	1.8	785	6	US-08-483-857-3	Sequence 4, Appl
C 126	1.9	34303	3	US-09-244-752-4	Sequence 4, Appl	199	15	1.8	789	3	US-09-107-532A-1489	Sequence 1489, Ap
C 127	1.9	34303	3	US-09-562-919-4	Sequence 4, Appl	200	15	1.8	834	4	US-09-280-590A-27	Sequence 27, Appl
C 128	1.9	34382	2	US-08-374-483-6	Sequence 6, Appl	201	15	1.8	850	3	US-08-928-941D-27	Sequence 27, Appl
C 129	1.9	34382	2	US-08-973-334-3	Sequence 3, Appl	202	15	1.8	850	4	US-09-280-590A-27	Sequence 27, Appl
C 130	1.9	35408	3	US-09-563-869A-3	Sequence 3, Appl	203	15	1.8	900	4	US-09-134-001C-2390	Sequence 2390, Ap
C 131	1.9	35408	3	US-08-549-489-3	Sequence 3, Appl	204	15	1.8	906	4	US-09-205-258-190	Sequence 190, App
C 132	1.9	35935	2	US-08-735-609-1	Sequence 1, Appl	205	15	1.8	913	1	US-08-109-391A-1	Sequence 1, Appl
C 133	1.9	35935	2	US-08-735-609-1	Sequence 1, Appl	206	15	1.8	913	2	US-08-459-019A-1	Sequence 1, Appl
C 134	1.9	35935	2	US-08-379-452-43	Sequence 43, Appl	207	15	1.8	913	2	US-08-459-019A-1	Sequence 1, Appl
C 135	1.9	35935	3	US-09-315-372-1	Sequence 1, Appl	208	15	1.8	913	3	US-08-458-860A-1	Sequence 1, Appl
C 136	1.9	35935	3	US-09-244-752-1	Sequence 1, Appl	209	15	1.8	936	3	US-09-175-172-1	Sequence 1, Appl
C 137	1.9	35935	3	US-09-244-752-1	Sequence 1, Appl	210	15	1.8	936	3	US-09-175-172-1	Sequence 1, Appl
C 138	1.9	35935	3	US-09-409-670-43	Sequence 43, Appl	211	15	1.8	945	3	US-08-682-080-10	Sequence 10, Appl
C 139	1.9	35935	4	US-09-562-919-1	Sequence 18, Appl	212	15	1.8	945	3	US-09-328-352-1198	Sequence 1198, Ap
C 140	1.9	40000	4	US-09-780-049-18	Sequence 18, Appl	213	15	1.8	966	2	US-08-766-738-2	Sequence 2, Appl
C 141	1.9	51952	3	US-08-947-823-1	Sequence 1, Appl	214	15	1.8	966	4	US-09-262-610-2	Sequence 2, Appl
C 142	1.9	80246	3	US-09-078-294-4	Sequence 4, Appl	215	15	1.8	972	4	US-09-134-001C-952	Sequence 952, App
C 143	1.9	80595	4	US-09-078-294-3	Sequence 3, Appl	216	15	1.8	1032	4	US-09-016-334-1338	Sequence 1338, Ap
C 144	1.9	90541	4	US-09-759-359A-3	Sequence 3, Appl	217	15	1.8	1065	4	US-09-634-238-121	Sequence 121, App
C 145	1.9	586073	4	US-08-545-528D-1	Sequence 1, Appl	218	15	1.8	1073	3	US-08-960-780-35	Sequence 35, Appl
C 146	1.9	586073	5	PCT-US91-02766-8	Sequence 8, Appl	219	15	1.8	1073	3	US-08-960-780-37	Sequence 37, Appl
C 147	1.8	25	1	US-08-447-411-79	Sequence 79, Appl	220	15	1.8	1073	3	US-09-073-898-35	Sequence 35, Appl
C 148	1.8	27	2	US-08-859-998-1032	Sequence 1032, Ap	221	15	1.8	1073	3	US-09-073-898-37	Sequence 37, Appl
C 149	1.8	27	4	US-09-225-928-1032	Sequence 1032, Ap	222	15	1.8	1073	4	US-09-252-991A-14969	Sequence 14969, A
C 150	1.8	27	4	US-09-225-928-1032	Sequence 1032, Ap	223	15	1.8	1076	3	US-08-960-780-18	Sequence 18, Appl
C 151	1.8	35	1	US-07-844-297-2	Sequence 2, Appl	224	15	1.8	1076	3	US-09-073-898-18	Sequence 18, Appl
C 152	1.8	35	1	US-07-844-297-2	Sequence 2, Appl	225	15	1.8	1077	4	US-09-328-352-2249	Sequence 2249, Ap
C 153	1.8	51	3	US-09-461-697-355	Sequence 355, App	226	15	1.8	1125	4	US-09-107-532A-2333	Sequence 2333, Ap
C 154	1.8	57	3	US-09-461-697-357	Sequence 357, App	227	15	1.8	1131	4	US-09-107-532A-2794	Sequence 2794, App
C 155	1.8	74	3	US-08-472-761-1	Sequence 1, Appl	228	15	1.8	1152	4	US-09-252-991A-14781	Sequence 14781, A
C 156	1.8	195	4	US-09-107-532A-996	Sequence 996, App	229	15	1.8	1158	4	US-08-545-593A-4	Sequence 4, Appl
C 157	1.8	199	4	US-09-312-294A-3083	Sequence 3083, Ap	230	15	1.8	1194	4	US-09-230-132-31	Sequence 31, Appl
C 158	1.8	235	4	US-08-675-499A-23	Sequence 23, Appl	231	15	1.8	1196	2	US-08-757-046A-1	Sequence 1, Appl
C 159	1.8	269	4	US-09-016-434-194	Sequence 194, App	232	15	1.8	1196	3	US-09-447-208-1	Sequence 1, Appl
C 160	1.8	270	4	US-09-313-294A-2439	Sequence 2439, Ap	233	15	1.8	1196	3	US-09-135-988-1	Sequence 1, Appl
C 161	1.8	278	3	US-09-376-781-10	Sequence 10, Appl	234	15	1.8	1196	3	US-09-330-317B-1	Sequence 1, Appl
C 162	1.8	282	4	US-09-313-294A-651	Sequence 651, App	235	15	1.8	1196	3	US-09-277-716-1	Sequence 1, Appl
C 163	1.8	291	4	US-09-313-294A-6219	Sequence 6219, Ap	236	15	1.8	1196	3	US-08-597-374A-1	Sequence 1, Appl
C 164	1.8	300	4	US-09-313-294A-6392	Sequence 6392, Ap	237	15	1.8	1196	4	US-08-968-909-1	Sequence 1, Appl
C 165	1.8	308	4	US-09-313-294A-5878	Sequence 5878, Ap	238	15	1.8	1196	4	US-09-609-161B-1	Sequence 1, Appl
C 166	1.8	311	4	US-09-702-705-855	Sequence 855, App	239	15	1.8	1196	4	US-09-808-589A-1	Sequence 1, Appl
C 167	1.8	311	4	US-09-736-457-855	Sequence 855, App	240	15	1.8	1215	4	US-08-990-103-1	Sequence 1, Appl
C 168	1.8	339	1	US-08-207-169A-1	Sequence 1, Appl	241	15	1.8	1215	4	US-09-205-258-59	Sequence 59, Appl
C 169	1.8	345	3	US-08-991-789A-209	Sequence 209, App	242	15	1.8	1217	3	US-09-277-716-17	Sequence 17, Appl
C 170	1.8	345	4	US-09-062-451-209	Sequence 209, App	243	15	1.8	1217	4	US-09-609-161B-17	Sequence 17, Appl
C 171	1.8	345	4	US-09-598-326-209	Sequence 209, App	244	15	1.8	1245	4	US-09-107-532A-3403	Sequence 3403, Ap
C 172	1.8	345	4	US-09-289-198-209	Sequence 209, App	245	15	1.8	1250	1	US-08-117-083-59	Sequence 59, Appl
C 173	1.8	366	3	US-08-908-643C-54	Sequence 54, Appl	246	15	1.8	1263	4	US-08-936-165A-39	Sequence 39, Appl

247	15	1.8	1271	3	US-09-330-317B-3	Sequence 3, Appl1	C 320	15	1.8	2061	3	US-08-751-230-9	Sequence 9, Appl1
248	15	1.8	1271	3	US-09-330-317B-11	Sequence 11, Appl1	C 321	15	1.8	2061	3	US-09-499-082-9	Sequence 9, Appl1
249	15	1.8	1271	3	US-09-330-317B-13	Sequence 13, Appl1	C 322	15	1.8	2061	3	US-09-258-372-9	Sequence 9, Appl1
250	15	1.8	1271	3	US-09-330-317B-15	Sequence 15, Appl1	C 323	15	1.8	2061	3	US-09-006-783A-2	Sequence 2, Appl1
251	15	1.8	1271	4	US-09-808-589A-3	Sequence 3, Appl1	C 324	15	1.8	2061	4	US-09-159-871-1	Sequence 1, Appl1
252	15	1.8	1271	4	US-09-808-589A-11	Sequence 11, Appl1	C 325	15	1.8	2065	4	US-09-129-666-5	Sequence 5, Appl1
253	15	1.8	1271	4	US-09-808-589A-13	Sequence 13, Appl1	C 326	15	1.8	2124	1	US-08-803-973-11	Sequence 11, Appl1
254	15	1.8	1271	4	US-09-808-589A-15	Sequence 15, Appl1	C 327	15	1.8	2124	1	US-08-803-973-11	Sequence 11, Appl1
C 255	15	1.8	1287	1	US-09-475-316A-67	Sequence 67, Appl1	C 328	15	1.8	2135	3	US-08-430-268A-1	Sequence 1, Appl1
256	15	1.8	1327	1	US-08-507-431-3	Sequence 3, Appl1	C 329	15	1.8	2156	3	US-08-965-762-16	Sequence 16, Appl1
257	15	1.8	1327	2	US-08-700-546-1	Sequence 1, Appl1	C 330	15	1.8	2156	4	US-09-911-927-16	Sequence 16, Appl1
258	15	1.8	1327	2	US-08-902-655A-3	Sequence 3, Appl1	C 331	15	1.8	2156	4	US-09-911-927-16	Sequence 16, Appl1
259	15	1.8	1327	2	US-09-116-622-3	Sequence 3, Appl1	C 332	15	1.8	2156	4	US-09-911-882-16	Sequence 16, Appl1
260	15	1.8	1327	3	US-09-219-277-3	Sequence 3, Appl1	C 333	15	1.8	2156	4	US-09-911-882-16	Sequence 16, Appl1
261	15	1.8	1327	3	US-09-599-661-3	Sequence 3, Appl1	C 334	15	1.8	2156	4	US-09-911-888-16	Sequence 16, Appl1
C 262	15	1.8	1328	4	US-09-482-273-84	Sequence 84, Appl1	C 335	15	1.8	2156	4	US-09-911-888-16	Sequence 16, Appl1
263	15	1.8	1338	4	US-09-328-352-1988	Sequence 1988, Ap	C 336	15	1.8	2163	4	US-09-186-276B-1	Sequence 1, Appl1
C 264	15	1.8	1344	4	US-09-328-352-3257	Sequence 3257, Ap	C 337	15	1.8	2163	4	US-08-842-443-1	Sequence 1, Appl1
C 265	15	1.8	1420	4	US-09-221-017B-847	Sequence 847, Ap	C 338	15	1.8	2163	4	US-09-186-188B-1	Sequence 1, Appl1
C 266	15	1.8	1431	4	US-09-134-001C-1280	Sequence 1280, Ap	C 339	15	1.8	2187	4	US-09-134-001C-2131	Sequence 2131, Ap
267	15	1.8	1435	2	US-08-878-989-14	Sequence 14, Appl1	C 340	15	1.8	2265	3	US-08-906-865-2	Sequence 2, Appl1
268	15	1.8	1435	2	US-09-272-796-14	Sequence 14, Appl1	C 341	15	1.8	2265	4	US-09-129-666-2	Sequence 2, Appl1
C 269	15	1.8	1463	4	US-09-393-634-6	Sequence 6, Appl1	C 342	15	1.8	2289	4	US-09-107-532A-621	Sequence 621, Ap
C 270	15	1.8	1464	4	US-09-328-352-3361	Sequence 3361, Ap	C 343	15	1.8	2291	1	US-08-920-812-9	Sequence 9, Appl1
C 271	15	1.8	1490	1	US-08-471-601-1	Sequence 1, Appl1	C 344	15	1.8	2291	1	US-08-920-823-9	Sequence 9, Appl1
C 272	15	1.8	1490	1	US-08-474-556-1	Sequence 1, Appl1	C 345	15	1.8	2291	1	US-08-921-177-9	Sequence 9, Appl1
C 273	15	1.8	1490	1	US-08-351-899-1	Sequence 1, Appl1	C 346	15	1.8	2291	1	US-08-362-577C-9	Sequence 9, Appl1
C 274	15	1.8	1490	1	US-08-479-382-1	Sequence 1, Appl1	C 347	15	1.8	2291	2	US-08-920-828-9	Sequence 9, Appl1
C 275	15	1.8	1490	1	US-08-470-354-1	Sequence 1, Appl1	C 348	15	1.8	2297	4	US-09-591-095-19	Sequence 19, Appl1
C 276	15	1.8	1490	1	US-08-479-383-1	Sequence 1, Appl1	C 349	15	1.8	2308	3	US-09-882-256-9	Sequence 9, Appl1
C 277	15	1.8	1490	2	US-08-479-041-1	Sequence 1, Appl1	C 350	15	1.8	2308	3	US-09-395-115-9	Sequence 9, Appl1
C 278	15	1.8	1490	2	US-08-819-646-1	Sequence 1, Appl1	C 351	15	1.8	2308	4	US-08-436-265-9	Sequence 9, Appl1
C 279	15	1.8	1490	4	US-09-195-716-1	Sequence 1, Appl1	C 352	15	1.8	2308	4	US-09-679-187-9	Sequence 9, Appl1
280	15	1.8	1491	4	US-09-252-991A-14836	Sequence 14836, A	C 353	15	1.8	2358	4	US-09-134-001C-1029	Sequence 1029, Ap
281	15	1.8	1501	3	US-09-196-520-1	Sequence 1, Appl1	C 354	15	1.8	2360	3	US-08-916-043-3	Sequence 3, Appl1
C 282	15	1.8	1503	4	US-09-482-273-17	Sequence 17, Appl1	C 355	15	1.8	2360	4	US-09-428-929-3	Sequence 3, Appl1
C 283	15	1.8	1520	4	US-09-620-312D-458	Sequence 458, Ap	C 356	15	1.8	2417	4	US-08-976-255-4	Sequence 4, Appl1
C 284	15	1.8	1529	4	US-09-149-476-2501	Sequence 250, Ap	C 357	15	1.8	2424	3	US-09-234-393-1	Sequence 1, Appl1
285	15	1.8	1548	4	US-09-453-702B-86	Sequence 86, Ap	C 358	15	1.8	2424	4	US-09-360-545-15	Sequence 15, Appl1
286	15	1.8	1629	5	PCT-US91-02560-3	Sequence 3, Appl1	C 359	15	1.8	2424	4	US-09-865-171-1	Sequence 1, Appl1
287	15	1.8	1636	3	US-08-908-643C-1	Sequence 1, Appl1	C 360	15	1.8	2424	4	US-09-398-395A-45	Sequence 45, Appl1
C 288	15	1.8	1650	1	US-08-399-561-5	Sequence 5, Appl1	C 361	15	1.8	2424	4	US-09-887-586A-45	Sequence 45, Appl1
C 289	15	1.8	1654	4	US-09-634-238-16	Sequence 16, Appl1	C 362	15	1.8	2424	4	US-09-895-752-45	Sequence 45, Appl1
C 290	15	1.8	1655	4	US-08-771-850A-1	Sequence 1, Appl1	C 363	15	1.8	2424	4	US-09-803-012B-45	Sequence 45, Appl1
C 291	15	1.8	1676	3	US-09-009-443-1	Sequence 1, Appl1	C 364	15	1.8	2496	4	US-09-269-861A-7	Sequence 7, Appl1
292	15	1.8	1677	2	US-07-868-353A-11	Sequence 11, Appl1	C 365	15	1.8	2503	3	US-09-149-476-249	Sequence 249, Ap
293	15	1.8	1703	3	US-07-868-353A-11	Sequence 11, Appl1	C 366	15	1.8	2525	4	US-09-234-393-39	Sequence 39, Appl1
294	15	1.8	1703	3	US-08-407-804-20	Sequence 20, Appl1	C 367	15	1.8	2525	4	US-09-865-171-39	Sequence 39, Appl1
295	15	1.8	1703	3	US-09-124-807-20	Sequence 20, Appl1	C 368	15	1.8	2528	3	US-09-234-393-37	Sequence 37, Appl1
296	15	1.8	1725	4	US-09-134-001C-545	Sequence 545, Ap	C 369	15	1.8	2528	3	US-09-234-393-41	Sequence 41, Appl1
C 297	15	1.8	1731	4	US-09-328-352-3287	Sequence 3287, Ap	C 370	15	1.8	2528	4	US-09-865-171-37	Sequence 37, Appl1
C 298	15	1.8	1781	4	US-09-040-229B-3	Sequence 3, Appl1	C 371	15	1.8	2528	4	US-09-865-171-41	Sequence 41, Appl1
C 299	15	1.8	1800	3	US-08-436-748-1	Sequence 1, Appl1	C 372	15	1.8	2542	3	US-09-330-317B-9	Sequence 9, Appl1
C 300	15	1.8	1800	3	US-08-483-857-1	Sequence 1, Appl1	C 373	15	1.8	2542	3	US-09-808-589A-9	Sequence 9, Appl1
C 301	15	1.8	1848	4	US-09-205-258-188	Sequence 188, Ap	C 374	15	1.8	2571	3	US-09-234-393-12	Sequence 12, Appl1
C 302	15	1.8	1861	4	US-09-129-668-7	Sequence 7, Appl1	C 375	15	1.8	2571	4	US-09-865-171-11	Sequence 11, Appl1
C 303	15	1.8	1863	4	US-08-545-573A-7	Sequence 7, Appl1	C 376	15	1.8	2736	4	US-09-220-132-66	Sequence 66, Appl1
304	15	1.8	1875	4	US-09-620-312D-455	Sequence 455, Ap	C 377	15	1.8	2778	1	US-08-702-054-1	Sequence 1, Appl1
305	15	1.8	1881	4	US-08-545-573A-8	Sequence 8, Appl1	C 378	15	1.8	2778	1	US-08-446-923-1	Sequence 1, Appl1
C 306	15	1.8	1902	2	US-09-258-257-1	Sequence 1, Appl1	C 379	15	1.8	2841	4	US-09-328-352-3847	Sequence 3847, Ap
C 307	15	1.8	1902	2	US-09-258-371-1	Sequence 1, Appl1	C 380	15	1.8	2861	4	US-09-016-433-1103	Sequence 1103, Ap
C 308	15	1.8	1902	3	US-08-569-721A-1	Sequence 1, Appl1	C 381	15	1.8	2934	1	US-08-178-446B-16	Sequence 16, Appl1
C 309	15	1.8	1902	3	US-08-751-230-1	Sequence 1, Appl1	C 382	15	1.8	2934	2	US-08-870-693-16	Sequence 16, Appl1
C 310	15	1.8	1902	3	US-09-499-082-1	Sequence 1, Appl1	C 383	15	1.8	2965	2	US-08-460-570-1	Sequence 1, Appl1
C 311	15	1.8	1902	3	US-09-258-372-1	Sequence 1, Appl1	C 384	15	1.8	2965	2	US-08-460-570-2	Sequence 2, Appl1
C 312	15	1.8	1965	3	US-09-461-697-292	Sequence 292, Ap	C 385	15	1.8	2965	3	US-08-286-870A-1	Sequence 1, Appl1
313	15	1.8	1975	1	US-08-484-105-11	Sequence 11, Appl1	C 386	15	1.8	3065	3	US-08-286-870A-2	Sequence 2, Appl1
314	15	1.8	1975	1	US-08-484-105-11	Sequence 11, Appl1	C 387	15	1.8	3065	4	US-09-539-333D-129	Sequence 129, Ap
315	15	1.8	2022	1	US-08-803-973-6	Sequence 6, Appl1	C 388	15	1.8	3088	1	US-08-418-444A-1	Sequence 1, Appl1
316	15	1.8	2022	1	US-08-803-972-6	Sequence 6, Appl1	C 389	15	1.8	3143	1	US-08-485-621-1	Sequence 1, Appl1
C 317	15	1.8	2036	4	US-09-702-705-1799	Sequence 1799, Ap	C 390	15	1.8	3143	2	US-08-973-831-1	Sequence 1, Appl1
C 318	15	1.8	2036	4	US-09-726-457-1799	Sequence 1799, Ap	C 391	15	1.8	3143	5	PCT-US96-09550A-1	Sequence 1, Appl1
C 319	15	1.8	2061	2	US-09-258-371-9	Sequence 9, Appl1	C 392	15	1.8	3169	2	US-08-588-258B-27	Sequence 27, Appl1

C 393	15	1.8	3169	3	US-08-460-505-27	Sequence 27, Appl	466	15	1.8	5948	4	US-09-017-947-1	Sequence 1, Appl
C 394	15	1.8	3169	5	PCT-US96-08295-27	Sequence 27, Appl	C 467	15	1.8	6030	1	US-08-441-139-8	Sequence 8, Appl
C 395	15	1.8	3170	3	US-09-344-520-1	Sequence 1, Appl	C 468	15	1.8	6133	4	US-09-453-7038-15	Sequence 15, Appl
C 396	15	1.8	3192	3	US-09-259-706-2	Sequence 2, Appl	469	15	1.8	6202	1	US-08-484-101B-41	Sequence 41, Appl
C 397	15	1.8	3192	4	US-09-388-995-2	Sequence 7, Appl	470	15	1.8	6202	3	US-08-714-524D-41	Sequence 5, Appl
398	15	1.8	3357	3	US-08-726-214-7	Sequence 5, Appl	471	15	1.8	6216	3	US-09-415-522-5	Sequence 126, App
399	15	1.8	3466	1	US-08-551-459-5	Sequence 8, Appl	C 472	15	1.8	6314	3	US-08-927-219-126	Sequence 98, Appl
400	15	1.8	3603	3	US-08-908-643C-85	Sequence 113, App	C 473	15	1.8	7100	4	US-09-620-3120-98	Sequence 2, Appl
401	15	1.8	3628	3	US-08-480-640A-113	Sequence 113, App	474	15	1.8	7494	4	US-09-470-661A-2	Sequence 104B, Ap
402	15	1.8	3628	3	US-08-295-802-113	Sequence 113, App	475	15	1.8	8312	4	US-09-620-3120-1048	Sequence 16, Appl
403	15	1.8	3628	3	US-08-488-237A-113	Sequence 113, App	C 477	15	1.8	8411	4	US-08-961-527-16	Sequence 25, Appl
404	15	1.8	3628	4	US-08-375-992A-113	Sequence 113, App	C 478	15	1.8	8900	4	US-09-505-984-25	Sequence 11, Appl
405	15	1.8	3628	4	US-08-472-679H-113	Sequence 89, Appl	C 479	15	1.8	8900	4	US-09-073-541A-25	Sequence 3, Appl
406	15	1.8	3645	4	US-08-908-643C-83	Sequence 11, Appl	C 480	15	1.8	9046	5	PCT-US95-0468B-1	Sequence 29, Appl
C 407	15	1.8	3668	4	US-09-302-620B-89	Sequence 8, Appl	C 481	15	1.8	10254	4	US-08-661-527-29	Sequence 2, Appl
C 408	15	1.8	3668	4	US-09-912-161-11	Sequence 11, Appl	C 482	15	1.8	10254	4	US-08-661-527-29	Sequence 2, Appl
409	15	1.8	3705	3	US-09-282-996-3	Sequence 84, Appl	C 483	15	1.8	10968	2	US-09-620-3120-1048	Sequence 1, Appl
410	15	1.8	3745	3	US-08-908-643C-84	Sequence 84, Appl	484	15	1.8	10968	2	US-09-620-3120-1048	Sequence 1, Appl
411	15	1.8	3770	4	US-09-221-017B-807	Sequence 807, Appl	485	15	1.8	12949	3	US-09-428-246-1	Sequence 1, Appl
412	15	1.8	3787	3	US-08-908-643C-82	Sequence 807, Appl	486	15	1.8	12949	3	US-09-428-246-1	Sequence 1, Appl
C 413	15	1.8	3900	4	US-09-302-620B-88	Sequence 807, Appl	487	15	1.8	13146	2	US-08-724-354D-13	Sequence 11, Appl
C 414	15	1.8	3900	4	US-09-302-620B-88	Sequence 807, Appl	487	15	1.8	13146	2	US-08-724-354D-13	Sequence 11, Appl
C 415	15	1.8	3909	4	US-08-232-537-1	Sequence 10, Appl	488	15	1.8	13146	3	US-09-370-984B-3	Sequence 3, Appl
C 416	15	1.8	3937	3	US-08-586-165-8	Sequence 8, Appl	489	15	1.8	15450	4	US-09-470-661A-1	Sequence 1, Appl
417	15	1.8	3942	3	US-08-480-640A-189	Sequence 189, App	C 490	15	1.8	16535	4	US-08-961-527-74	Sequence 74, Appl
418	15	1.8	3942	3	US-08-488-237A-189	Sequence 189, App	C 491	15	1.8	16836	4	US-09-147-236-1	Sequence 1, Appl
419	15	1.8	3942	3	US-08-488-237A-189	Sequence 189, App	C 492	15	1.8	16836	4	US-09-147-236-1	Sequence 1, Appl
420	15	1.8	3942	4	US-08-375-992A-189	Sequence 189, App	C 493	15	1.8	16836	4	US-09-522-474-1	Sequence 10, Appl
421	15	1.8	3942	4	US-08-472-679H-189	Sequence 189, App	C 494	15	1.8	16836	4	US-09-522-474-1	Sequence 10, Appl
422	15	1.8	3951	3	US-09-138-103-1	Sequence 1, Appl	C 495	15	1.8	19718	4	US-08-961-527-99	Sequence 99, Appl
423	15	1.8	3951	3	US-09-962-665-3	Sequence 3, Appl	496	15	1.8	28882	3	US-08-961-527-140	Sequence 140, Appl
424	15	1.8	3957	2	US-08-304-309-1	Sequence 1, Appl	C 497	15	1.8	43676	4	US-09-356-952-12	Sequence 12, Appl
425	15	1.8	3957	3	US-08-991-942-1	Sequence 1, Appl	C 498	15	1.8	45716	4	US-08-965-048-5	Sequence 5, Appl
C 426	15	1.8	3997	4	US-09-409-648-2	Sequence 2, Appl	C 499	15	1.8	45716	4	US-08-965-048-5	Sequence 5, Appl
C 427	15	1.8	3997	4	US-09-409-648-2	Sequence 2, Appl	C 500	15	1.8	46718	4	US-09-816-093-3	Sequence 3, Appl
C 428	15	1.8	4066	3	US-09-215-966-51	Sequence 21, Appl	501	15	1.8	64457	4	US-09-803-671B-3	Sequence 3, Appl
429	15	1.8	4079	4	US-09-016-434-1222	Sequence 1222, Ap	C 502	15	1.8	814857	4	US-09-801-876B-3	Sequence 1, Appl
430	15	1.8	4079	4	US-09-016-434-1247	Sequence 1247, Ap	503	15	1.8	814857	4	US-09-801-876B-3	Sequence 1, Appl
C 431	15	1.8	4098	4	US-09-268-866-1	Sequence 1, Appl	504	15	1.8	819608	4	US-09-539-333D-1	Sequence 1, Appl
432	15	1.8	4138	1	US-08-447-411-75	Sequence 75, Appl	505	15	1.8	8319608	4	US-08-545-528D-1	Sequence 1, Appl
433	15	1.8	4138	2	US-08-662-227-33	Sequence 33, Appl	506	15	1.8	8580073	4	US-09-198-452A-1	Sequence 1, Appl
434	15	1.8	4138	4	US-09-017-947-93	Sequence 33, Appl	507	15	1.8	1230025	4	US-09-198-452A-1	Sequence 1, Appl
435	15	1.8	4368	5	PCT-US95-04567-3	Sequence 3, Appl	C 508	15	1.8	1230025	4	US-09-198-452A-1	Sequence 1, Appl
436	15	1.8	4371	1	US-08-803-973-1	Sequence 1, Appl	509	15	1.8	1830121	4	US-09-557-88A-1	Sequence 1, Appl
437	15	1.8	4371	1	US-08-803-972-1	Sequence 1, Appl	510	15	1.8	1830121	4	US-09-557-88A-1	Sequence 1, Appl
438	15	1.8	4407	4	US-08-976-259-75	Sequence 75, Appl	511	15	1.8	4403765	3	US-09-103-840A-2	Sequence 2, Appl
439	15	1.8	4515	4	US-09-620-312D-146	Sequence 146, App	512	15	1.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl
C 440	15	1.8	4673	5	PCT-US92-00018-1	Sequence 1, Appl	C 513	14	1.7	17	2	US-08-292-620A-1722	Sequence 1722, Ap
C 441	15	1.8	4856	3	US-09-045-360-1	Sequence 1, Appl	C 514	14	1.7	17	3	US-09-071-845-1722	Sequence 1722, Ap
C 442	15	1.8	5000	4	US-08-836-687B-35	Sequence 35, Appl	C 515	14	1.7	18	4	US-09-422-978-11087	Sequence 51102, A
443	15	1.8	5405	3	US-09-282-996-1	Sequence 1, Appl	C 516	14	1.7	19	4	US-09-422-978-5412	Sequence 1417, Ap
444	15	1.8	5427	3	US-09-282-996-2	Sequence 2, Appl	C 517	14	1.7	20	4	US-09-198-452A-26	Sequence 26, Appl
445	15	1.8	5544	3	US-08-851-843A-68	Sequence 68, Appl	C 518	14	1.7	21	1	US-08-474-633A-48	Sequence 48, Appl
446	15	1.8	5544	3	US-08-974-549A-111	Sequence 111, App	C 519	14	1.7	21	1	US-08-823-771-48	Sequence 26, Appl
447	15	1.8	5544	3	US-08-854-050-68	Sequence 68, Appl	C 520	14	1.7	21	5	PCT-US92-06411-26	Sequence 15, Appl
448	15	1.8	5544	4	US-09-430-323-68	Sequence 111, App	C 521	14	1.7	24	2	US-08-819-912-15	Sequence 33, Appl
449	15	1.8	5544	4	US-08-912-951-111	Sequence 111, App	C 522	14	1.7	24	2	US-08-577-858A-33	Sequence 21, Appl
C 450	15	1.8	5558	4	US-08-961-527-103	Sequence 103, App	C 523	14	1.7	24	3	US-09-009-483A-21	Sequence 6, Appl
C 451	15	1.8	5596	4	US-09-911-927-1	Sequence 1, Appl	C 524	14	1.7	24	3	US-09-311-260-6	Sequence 5, Appl
452	15	1.8	5596	4	US-09-911-927-1	Sequence 1, Appl	C 525	14	1.7	30	5	PCT-US93-09695-5	Sequence 11, Appl
453	15	1.8	5596	4	US-09-911-927-3	Sequence 3, Appl	526	14	1.7	30	5	PCT-US93-09695-5	Sequence 11, Appl
C 454	15	1.8	5596	4	US-09-911-927-3	Sequence 3, Appl	C 527	14	1.7	31	3	US-08-841-483-11	Sequence 8, Appl
C 455	15	1.8	5596	4	US-09-911-927-3	Sequence 3, Appl	C 528	14	1.7	31	3	US-08-841-483-11	Sequence 8, Appl
C 456	15	1.8	5596	4	US-09-911-927-3	Sequence 3, Appl	C 529	14	1.7	31	3	US-08-841-483-11	Sequence 8, Appl
457	15	1.8	5596	4	US-09-911-927-3	Sequence 3, Appl	C 530	14	1.7	31	3	US-08-841-483-11	Sequence 8, Appl
C 458	15	1.8	5596	4	US-09-911-927-3	Sequence 3, Appl	C 531	14	1.7	31	3	US-08-841-483-11	Sequence 8, Appl
C 459	15	1.8	5733	2	US-08-473-553A-1	Sequence 1, Appl	C 532	14	1.7	39	3	US-08-444-644-8	Sequence 904, App
460	15	1.8	5785	3	US-08-480-640A-221	Sequence 221, App	C 533	14	1.7	47	4	US-09-671-317-904	Sequence 57, Appl
461	15	1.8	5785	3	US-08-686-968C-221	Sequence 221, App	C 534	14	1.7	49	2	US-08-484-993B-57	Sequence 57, Appl
462	15	1.8	5785	3	US-08-488-237A-221	Sequence 221, App	C 535	14	1.7	49	2	US-08-484-993B-57	Sequence 57, Appl
463	15	1.8	5785	4	US-08-472-679H-221	Sequence 221, App	C 536	14	1.7	49	2	US-08-484-993B-57	Sequence 57, Appl
464	15	1.8	5924	1	US-08-447-411-44	Sequence 44, Appl	C 537	14	1.7	49	2	US-08-480-150A-57	Sequence 57, Appl
465	15	1.8	5948	2	US-08-662-227-1	Sequence 1, Appl	C 538	14	1.7	49	3	US-08-458-731-57	Sequence 57, Appl

C 539	14	1.7	49	3	US-08-149-223A-57	Sequence 57, Appl	612	14	1.7	350	4	US-09-107-532A-603	Sequence 603, App
C 540	14	1.7	72	4	US-09-423-439-45	Sequence 45, Appl	613	14	1.7	331	4	US-09-389-681-228	Sequence 228, App
C 541	14	1.7	72	4	US-09-423-439-45	Sequence 46, Appl	614	14	1.7	331	4	US-09-620-405B-228	Sequence 228, App
C 542	14	1.7	97	1	US-08-182-175A-54	Sequence 54, Appl	615	14	1.7	331	4	US-09-339-338-228	Sequence 228, App
C 543	14	1.7	97	1	US-08-474-633A-63	Sequence 63, Appl	616	14	1.7	331	4	US-09-433-826B-228	Sequence 228, App
C 544	14	1.7	97	1	US-08-823-771-63	Sequence 63, Appl	617	14	1.7	331	4	US-09-604-287A-228	Sequence 228, App
C 545	14	1.7	97	5	PCT-US92-06412-54	Sequence 54, Appl	618	14	1.7	338	3	US-09-385-982-524	Sequence 524, App
C 546	14	1.7	111	1	US-08-105-483-309	Sequence 309, App	619	14	1.7	339	4	US-09-328-352-1872	Sequence 1872, App
C 547	14	1.7	111	1	US-08-709-209-309	Sequence 309, App	620	14	1.7	406	3	US-09-386-493-13	Sequence 13, Appl
C 548	14	1.7	111	1	US-08-458-101-309	Sequence 309, App	621	14	1.7	414	4	US-08-976-259-2	Sequence 2, Appl
C 549	14	1.7	122	4	US-09-313-294A-1173	Sequence 1173, App	622	14	1.7	424	1	US-08-470-179-141	Sequence 141, App
C 550	14	1.7	174	5	PCT-US95-09334-1	Sequence 1, Appl	623	14	1.7	429	4	US-09-328-352-2494	Sequence 2494, App
C 551	14	1.7	182	4	US-09-702-705-1517	Sequence 1517, App	624	14	1.7	430	4	US-09-222-575-147	Sequence 147, App
C 552	14	1.7	182	4	US-09-736-457-1517	Sequence 1517, App	625	14	1.7	430	4	US-09-389-681-147	Sequence 147, App
C 553	14	1.7	186	4	US-09-107-532A-616	Sequence 616, App	626	14	1.7	430	4	US-09-620-405B-147	Sequence 147, App
C 554	14	1.7	189	4	US-09-328-352-3382	Sequence 3382, App	627	14	1.7	430	4	US-09-339-338-147	Sequence 147, App
C 555	14	1.7	207	4	US-09-328-352-3382	Sequence 3382, App	628	14	1.7	430	4	US-09-433-826B-147	Sequence 147, App
C 556	14	1.7	223	3	US-08-991-789A-57	Sequence 57, Appl	629	14	1.7	430	4	US-09-604-287A-147	Sequence 147, App
C 557	14	1.7	223	4	US-09-062-451-57	Sequence 57, Appl	630	14	1.7	439	4	US-09-222-575-149	Sequence 149, App
C 558	14	1.7	223	4	US-09-598-326-57	Sequence 57, Appl	631	14	1.7	439	4	US-09-389-681-149	Sequence 149, App
C 559	14	1.7	223	4	US-09-289-198-57	Sequence 57, Appl	632	14	1.7	439	4	US-09-620-405B-149	Sequence 149, App
C 560	14	1.7	224	4	US-09-016-434-212	Sequence 212, App	633	14	1.7	439	4	US-09-339-338B-149	Sequence 149, App
C 561	14	1.7	225	4	US-09-134-001C-61	Sequence 61, Appl	634	14	1.7	439	4	US-09-433-826B-149	Sequence 149, App
C 562	14	1.7	229	4	US-09-049-698-15	Sequence 15, Appl	635	14	1.7	439	4	US-09-604-287A-149	Sequence 149, App
C 563	14	1.7	237	4	US-09-107-532A-2741	Sequence 2741, App	636	14	1.7	441	4	US-09-252-991A-6383	Sequence 6383, App
C 564	14	1.7	238	4	US-09-495-050A-125	Sequence 125, App	637	14	1.7	450	4	US-09-328-352-311	Sequence 311, App
C 565	14	1.7	241	4	US-09-389-681-356	Sequence 356, App	638	14	1.7	459	4	US-09-328-352-3395	Sequence 3395, App
C 566	14	1.7	241	4	US-09-433-826B-356	Sequence 356, App	639	14	1.7	461	3	US-09-087-232A-11	Sequence 11, Appl
C 567	14	1.7	241	4	US-09-604-287A-356	Sequence 356, App	640	14	1.7	469	3	US-09-008-979A-2	Sequence 2, Appl
C 568	14	1.7	244	4	US-08-435-684A-67	Sequence 67, Appl	641	14	1.7	469	3	US-09-460-618-2	Sequence 2, Appl
C 569	14	1.7	244	2	US-08-934-877A-67	Sequence 67, Appl	642	14	1.7	469	3	US-09-310-235B-2	Sequence 2, Appl
C 570	14	1.7	244	3	US-08-871-678C-67	Sequence 67, Appl	643	14	1.7	474	4	US-09-107-532A-1146	Sequence 1146, App
C 571	14	1.7	244	3	US-08-871-678C-67	Sequence 67, Appl	644	14	1.7	484	4	US-09-481-049-10	Sequence 10, Appl
C 572	14	1.7	246	4	US-09-107-532A-322	Sequence 322, App	645	14	1.7	486	4	US-09-107-532A-188	Sequence 188, App
C 573	14	1.7	246	4	US-09-107-532A-328	Sequence 328, App	646	14	1.7	491	3	US-08-784-582-52	Sequence 52, Appl
C 574	14	1.7	246	4	US-09-107-532A-329	Sequence 329, App	647	14	1.7	492	4	US-09-615-192A-223	Sequence 223, App
C 575	14	1.7	246	4	US-09-107-532A-330	Sequence 330, App	648	14	1.7	493	4	US-09-702-705-750	Sequence 750, App
C 576	14	1.7	246	4	US-09-107-532A-331	Sequence 331, App	649	14	1.7	500	3	US-09-736-457-750	Sequence 750, App
C 577	14	1.7	246	4	US-09-107-532A-332	Sequence 332, App	650	14	1.7	503	3	US-09-330-611-26	Sequence 26, Appl
C 578	14	1.7	246	4	US-09-107-532A-333	Sequence 333, App	651	14	1.7	513	4	US-09-228-986-49	Sequence 49, Appl
C 579	14	1.7	246	4	US-09-107-532A-334	Sequence 334, App	652	14	1.7	513	4	US-09-702-705-314	Sequence 314, App
C 580	14	1.7	246	4	US-09-107-532A-335	Sequence 335, App	653	14	1.7	513	4	US-09-736-457-314	Sequence 314, App
C 581	14	1.7	246	4	US-09-107-532A-336	Sequence 336, App	654	14	1.7	519	2	US-08-503-226B-36	Sequence 36, Appl
C 582	14	1.7	249	4	US-09-107-532A-130	Sequence 130, App	655	14	1.7	519	3	US-08-721-458B-36	Sequence 36, Appl
C 583	14	1.7	252	4	US-09-107-532A-237	Sequence 237, App	656	14	1.7	522	4	US-09-702-705-688	Sequence 688, App
C 584	14	1.7	258	4	US-09-134-001C-1128	Sequence 1128, App	657	14	1.7	522	4	US-09-736-457-688	Sequence 688, App
C 585	14	1.7	258	4	US-09-345-882-21	Sequence 21, Appl	658	14	1.7	528	4	US-09-105-542A-12	Sequence 12, Appl
C 586	14	1.7	263	4	US-09-313-294A-3143	Sequence 3143, App	659	14	1.7	541	2	US-08-503-226B-35	Sequence 35, Appl
C 587	14	1.7	273	2	US-08-737-298-1	Sequence 1, Appl	660	14	1.7	541	3	US-08-721-458B-35	Sequence 35, Appl
C 588	14	1.7	278	1	US-08-248-474-106	Sequence 106, App	661	14	1.7	549	4	US-09-252-991A-11414	Sequence 11414, A
C 589	14	1.7	278	3	US-08-756-849-106	Sequence 106, App	662	14	1.7	558	4	US-09-280-116-185	Sequence 185, App
C 590	14	1.7	278	3	US-08-446-935-3	Sequence 3, Appl	663	14	1.7	579	4	US-09-449-285A-13	Sequence 13, Appl
C 591	14	1.7	279	4	US-09-313-294A-1267	Sequence 1267, App	664	14	1.7	582	3	US-09-449-285A-13	Sequence 13, Appl
C 592	14	1.7	281	4	US-09-313-294A-1935	Sequence 1935, App	665	14	1.7	585	4	US-08-896-164-55	Sequence 55, Appl
C 593	14	1.7	282	4	US-09-313-294A-3951	Sequence 3951, App	666	14	1.7	587	1	US-09-107-532A-3148	Sequence 3148, App
C 594	14	1.7	283	4	US-09-313-294A-3036	Sequence 3036, App	667	14	1.7	587	1	US-08-742-023-7	Sequence 7, Appl
C 595	14	1.7	285	4	US-09-313-294A-6348	Sequence 6348, App	668	14	1.7	587	1	US-08-968-505-7	Sequence 7, Appl
C 596	14	1.7	289	4	US-09-313-294A-2779	Sequence 2779, App	669	14	1.7	591	2	US-08-704-473-1	Sequence 1, Appl
C 597	14	1.7	294	4	US-09-107-532A-278	Sequence 278, App	670	14	1.7	595	4	US-09-394-455-42	Sequence 42, Appl
C 598	14	1.7	300	4	US-09-313-294A-5183	Sequence 5183, App	671	14	1.7	597	2	US-08-687-080-72	Sequence 72, Appl
C 599	14	1.7	301	4	US-09-049-698-14	Sequence 14, Appl	672	14	1.7	603	3	US-09-252-991A-6165	Sequence 6165, App
C 600	14	1.7	301	4	US-09-252-991A-15559	Sequence 15559, A	673	14	1.7	605	3	US-09-328-111-430	Sequence 430, App
C 601	14	1.7	312	4	US-09-134-001C-303	Sequence 303, App	674	14	1.7	606	3	US-09-328-111-132	Sequence 132, App
C 602	14	1.7	315	4	US-09-107-532A-1881	Sequence 1881, App	675	14	1.7	610	4	US-09-484-970B-100	Sequence 100, App
C 603	14	1.7	320	4	US-09-134-001C-686	Sequence 686, App	676	14	1.7	612	4	US-09-766-055A-3	Sequence 3, Appl
C 604	14	1.7	330	1	US-08-620-467A-3	Sequence 3, Appl	677	14	1.7	612	4	US-09-107-532A-1950	Sequence 1950, App
C 605	14	1.7	339	3	US-08-348-572-3	Sequence 3, Appl	678	14	1.7	616	2	US-08-630-822A-93	Sequence 93, Appl
C 606	14	1.7	339	3	US-09-041-090B-3	Sequence 3, Appl	679	14	1.7	616	2	US-09-005-069-93	Sequence 93, Appl
C 607	14	1.7	350	4	US-09-404-879A-353	Sequence 353, App	680	14	1.7	616	2	US-09-171-156A-42	Sequence 42, Appl
C 608	14	1.7	351	4	US-09-107-532A-1880	Sequence 1880, App	681	14	1.7	616	4	US-09-004-730A-42	Sequence 42, Appl
C 609	14	1.7	361	3	US-08-943-731-95	Sequence 1874, App	682	14	1.7	616	4	US-08-981-799A-42	Sequence 42, Appl
C 610	14	1.7	366	4	US-09-134-001C-1134	Sequence 1134, App	683	14	1.7	621	4	US-09-107-532A-2883	Sequence 2883, App
C 611	14	1.7	385	4	US-08-747-562-20	Sequence 20, Appl	684	14	1.7	630	4	US-09-328-352-4081	Sequence 4081, App

C 685	14	1.7	630	4	US-09-107-532A-1815	Sequence 1815, Ap	C 758	14	1.7	841	2	US-08-743-637B-34	Sequence 34, Appl
C 686	14	1.7	636	3	US-09-328-111-370	Sequence 370, App	C 759	14	1.7	841	3	US-08-526-840B-34	Sequence 34, Appl
C 687	14	1.7	639	4	US-08-592-126-88	Sequence 88, Appl	C 760	14	1.7	852	4	US-09-134-001C-2654	Sequence 2654, Ap
C 688	14	1.7	639	4	US-09-482-273-49	Sequence 48, Appl	C 761	14	1.7	852	4	US-09-328-352-905	Sequence 905, App
C 689	14	1.7	639	4	US-09-168-595-88	Sequence 88, Appl	C 762	14	1.7	858	4	US-09-328-352-2038	Sequence 2038, Ap
C 690	14	1.7	642	4	US-09-134-001C-1132	Sequence 1132, Ap	C 763	14	1.7	859	3	US-09-247-373B-47	Sequence 47, Appl
C 691	14	1.7	646	4	US-09-016-434-1385	Sequence 1385, Ap	C 764	14	1.7	867	2	US-08-961-858-2	Sequence 2, Appl
C 692	14	1.7	648	4	US-09-107-532A-1324	Sequence 1324, Ap	C 765	14	1.7	867	2	US-08-961-858-4	Sequence 4, Appl
C 693	14	1.7	660	4	US-09-220-132-62	Sequence 62, App	C 766	14	1.7	867	3	US-09-089-593-2	Sequence 2, Appl
C 694	14	1.7	663	4	US-09-252-991A-12371	Sequence 12371, A	C 767	14	1.7	867	3	US-09-089-593-4	Sequence 4, Appl
C 695	14	1.7	669	4	US-09-280-116-187	Sequence 187, App	C 768	14	1.7	867	3	US-08-950-925-3	Sequence 3, Appl
C 696	14	1.7	676	3	US-08-998-416-485	Sequence 485, App	C 769	14	1.7	867	4	US-09-134-001C-2126	Sequence 2126, Ap
C 697	14	1.7	676	4	US-09-016-434-1022	Sequence 1022, Ap	C 770	14	1.7	867	4	US-09-565-286-3	Sequence 3, Appl
C 698	14	1.7	678	1	US-07-991-667B-23	Sequence 23, Appl	C 771	14	1.7	878	4	US-09-620-312D-683	Sequence 683, App
C 699	14	1.7	678	1	US-08-107-755A-23	Sequence 23, Appl	C 772	14	1.7	879	4	US-09-252-991A-12179	Sequence 12179, A
C 700	14	1.7	678	2	US-08-544-332-23	Sequence 23, Appl	C 773	14	1.7	886	4	US-09-134-001C-2806	Sequence 2806, Ap
C 701	14	1.7	678	4	US-09-370-861A-23	Sequence 23, Appl	C 774	14	1.7	885	3	US-08-784-582-55	Sequence 55, Appl
C 702	14	1.7	684	4	US-09-252-991A-12610	Sequence 12610, A	C 775	14	1.7	897	3	US-09-668-680-5	Sequence 5, Appl
C 703	14	1.7	684	4	US-09-328-352-581	Sequence 581, App	C 776	14	1.7	897	1	US-08-664-596B-23	Sequence 23, Appl
C 704	14	1.7	699	2	US-08-869-674-1	Sequence 1, Appl	C 777	14	1.7	907	2	US-08-739-775-1	Sequence 1, Appl
C 705	14	1.7	699	3	US-09-213-010-1	Sequence 1, Appl	C 778	14	1.7	909	4	US-09-107-532A-1981	Sequence 1981, Ap
C 706	14	1.7	699	4	US-09-213-011-1	Sequence 1, Appl	C 779	14	1.7	918	4	US-09-328-352-2919	Sequence 2919, Ap
C 707	14	1.7	703	4	US-09-669-751-64	Sequence 64, Appl	C 780	14	1.7	920	4	US-09-221-017B-556	Sequence 556, App
C 708	14	1.7	705	4	US-09-107-532A-599	Sequence 599, App	C 781	14	1.7	925	4	US-09-620-312D-963	Sequence 963, App
C 709	14	1.7	705	4	US-09-107-532A-600	Sequence 600, App	C 782	14	1.7	936	4	US-09-328-352-2429	Sequence 2429, Ap
C 710	14	1.7	705	4	US-09-107-532A-601	Sequence 601, App	C 783	14	1.7	948	4	US-09-446-250-1	Sequence 11432, A
C 711	14	1.7	705	4	US-09-107-532A-602	Sequence 602, App	C 784	14	1.7	948	4	US-09-393-634-57	Sequence 57, Appl
C 712	14	1.7	705	4	US-09-107-532A-604	Sequence 604, App	C 785	14	1.7	949	3	US-08-714-918-12	Sequence 12, Appl
C 713	14	1.7	705	4	US-09-107-532A-606	Sequence 606, App	C 786	14	1.7	949	3	US-09-265-315-12	Sequence 12, Appl
C 714	14	1.7	705	4	US-09-107-532A-607	Sequence 607, App	C 787	14	1.7	949	3	US-09-265-315-12	Sequence 12, Appl
C 715	14	1.7	705	4	US-09-107-532A-608	Sequence 608, App	C 788	14	1.7	949	3	US-09-266-417-12	Sequence 12, Appl
C 716	14	1.7	705	4	US-09-107-532A-619	Sequence 619, App	C 789	14	1.7	951	4	US-09-328-352-3474	Sequence 3474, Ap
C 717	14	1.7	707	5	US-08-591-468-25	Sequence 25, Appl	C 790	14	1.7	957	4	US-09-328-352-2743	Sequence 2743, Ap
C 718	14	1.7	707	5	BCT-US94-06430-25	Sequence 25, Appl	C 791	14	1.7	965	2	US-09-141-135-1	Sequence 1, Appl
C 719	14	1.7	708	4	US-09-107-532A-3325	Sequence 3325, Ap	C 792	14	1.7	966	4	US-09-252-991A-12102	Sequence 12102, A
C 720	14	1.7	710	3	US-09-231-529-8	Sequence 8, Appl	C 793	14	1.7	966	4	US-09-328-352-2143	Sequence 2143, Ap
C 721	14	1.7	710	3	US-08-977-816-8	Sequence 8, Appl	C 794	14	1.7	966	4	US-09-328-352-2550	Sequence 2550, Ap
C 722	14	1.7	717	2	US-08-963-743-3	Sequence 3, Appl	C 795	14	1.7	966	4	US-09-107-532A-2829	Sequence 2829, Ap
C 723	14	1.7	719	2	US-08-963-743-6	Sequence 6, Appl	C 796	14	1.7	969	4	US-09-042-775-1	Sequence 1, Appl
C 724	14	1.7	722	4	US-09-280-116-128	Sequence 128, App	C 797	14	1.7	978	1	US-08-446-486-31	Sequence 31, Appl
C 725	14	1.7	735	4	US-09-328-352-415	Sequence 415, App	C 798	14	1.7	978	1	US-08-463-308-31	Sequence 31, Appl
C 726	14	1.7	754	2	US-08-879-995A-2	Sequence 2, Appl	C 799	14	1.7	979	1	US-08-446-486-30	Sequence 30, Appl
C 727	14	1.7	754	3	US-09-215-096-2	Sequence 2, Appl	C 800	14	1.7	979	1	US-08-463-308-30	Sequence 30, Appl
C 728	14	1.7	755	3	US-09-174-437-39	Sequence 39, Appl	C 801	14	1.7	979	1	US-08-463-308-30	Sequence 30, Appl
C 729	14	1.7	755	4	US-09-686-055A-39	Sequence 39, Appl	C 802	14	1.7	987	6	5254799-30	Patent No. 5254799
C 730	14	1.7	756	4	US-09-107-532A-83	Sequence 83, Appl	C 803	14	1.7	987	4	US-09-328-352-410	Sequence 410, App
C 731	14	1.7	757	2	US-08-665-647-6	Sequence 6, Appl	C 804	14	1.7	990	2	US-08-410-167A-1	Sequence 1, Appl
C 732	14	1.7	757	3	US-08-998-416-454	Sequence 454, App	C 805	14	1.7	990	4	US-08-806-959-1	Sequence 1, Appl
C 733	14	1.7	764	4	US-08-858-207A-2	Sequence 2, Appl	C 806	14	1.7	993	1	US-09-205-258-132	Sequence 132, App
C 734	14	1.7	771	3	US-08-998-416-474	Sequence 474, App	C 807	14	1.7	993	1	US-08-705-377-1	Sequence 1, Appl
C 735	14	1.7	785	3	US-09-008-979A-6	Sequence 6, Appl	C 808	14	1.7	993	1	US-08-705-377-2	Sequence 2, Appl
C 736	14	1.7	785	3	US-09-460-618-6	Sequence 6, Appl	C 809	14	1.7	993	1	US-08-705-377-4	Sequence 4, Appl
C 737	14	1.7	785	3	US-09-310-235B-6	Sequence 6, Appl	C 810	14	1.7	993	1	US-08-705-377-5	Sequence 5, Appl
C 738	14	1.7	792	4	US-08-833-752-1	Sequence 1, Appl	C 811	14	1.7	993	1	US-08-705-377-6	Sequence 6, Appl
C 739	14	1.7	793	3	US-08-897-736-1	Sequence 1, Appl	C 812	14	1.7	993	2	US-09-052-962-1	Sequence 2, Appl
C 740	14	1.7	793	3	US-08-897-736-1	Sequence 1, Appl	C 813	14	1.7	993	2	US-09-052-962-2	Sequence 2, Appl
C 741	14	1.7	798	2	US-08-319-866-7	Sequence 7, Appl	C 814	14	1.7	993	2	US-09-052-962-3	Sequence 3, Appl
C 742	14	1.7	801	4	US-09-252-991A-7189	Sequence 7189, Ap	C 815	14	1.7	993	2	US-09-052-962-4	Sequence 4, Appl
C 743	14	1.7	801	4	US-09-252-991A-7458	Sequence 7458, Ap	C 816	14	1.7	993	2	US-09-052-962-5	Sequence 5, Appl
C 744	14	1.7	801	4	US-09-107-532A-3432	Sequence 3432, Ap	C 817	14	1.7	993	2	US-09-052-962-6	Sequence 6, Appl
C 745	14	1.7	804	4	US-09-328-352-3673	Sequence 3673, Ap	C 818	14	1.7	993	2	US-09-053-068-1	Sequence 1, Appl
C 746	14	1.7	807	4	US-09-107-532A-2143	Sequence 2143, Ap	C 819	14	1.7	993	2	US-09-053-068-2	Sequence 2, Appl
C 747	14	1.7	810	4	US-09-107-532A-2028	Sequence 2028, Ap	C 820	14	1.7	993	2	US-09-053-068-3	Sequence 3, Appl
C 748	14	1.7	813	4	US-09-328-352-2407	Sequence 2407, Ap	C 821	14	1.7	993	2	US-09-053-068-4	Sequence 4, Appl
C 749	14	1.7	813	4	US-09-107-532A-1566	Sequence 1566, Ap	C 822	14	1.7	993	2	US-09-053-068-5	Sequence 5, Appl
C 750	14	1.7	815	4	US-09-221-017B-962	Sequence 962, App	C 823	14	1.7	993	2	US-09-053-068-6	Sequence 6, Appl
C 751	14	1.7	816	4	US-09-328-475C-197	Sequence 1967, App	C 824	14	1.7	993	2	US-09-053-068-6	Sequence 6, Appl
C 752	14	1.7	825	4	US-09-134-001C-1569	Sequence 1569, Ap	C 825	14	1.7	993	2	US-09-101-126-2	Sequence 2, Appl
C 753	14	1.7	831	3	US-09-052-505-1	Sequence 1, Appl	C 826	14	1.7	1000	4	US-09-620-312D-962	Sequence 962, App
C 754	14	1.7	831	3	US-09-259-473A-1	Sequence 1, Appl	C 827	14	1.7	1001	4	US-08-882-501-32	Sequence 32, Appl
C 755	14	1.7	831	3	US-09-205-258-75	Sequence 75, Appl	C 828	14	1.7	1001	4	US-08-882-501-32	Sequence 32, Appl
C 756	14	1.7	838	4	US-09-221-017B-141	Sequence 141, App	C 829	14	1.7	1001	4	US-09-671-317-411	Sequence 411, App
C 757	14	1.7	840	3	US-09-042-771-1	Sequence 1, Appl	C 830	14	1.7	1001	4	US-09-671-317-419	Sequence 419, App



C 831	14	1.7	1002	4	US-09-641-638-591	Sequence 591, App	904	14	1.7	1248	4	US-09-134-001C-342	Sequence 342, App
C 832	14	1.7	1007	3	US-08-924-747-3	Sequence 3, Appl1	C 905	14	1.7	1248	5	PCT-US92-00331-2	Sequence 2, Appl1
C 833	14	1.7	1007	3	US-09-247-373B-3	Sequence 3, Appl1	C 906	14	1.7	1251	3	US-08-910-505-1	Sequence 1, Appl1
C 834	14	1.7	1007	3	US-09-296-715-3	Sequence 3, Appl1	C 907	14	1.7	1251	3	US-09-330-611-1	Sequence 1, Appl1
C 835	14	1.7	1008	4	US-09-394-455-3	Sequence 3, Appl1	C 908	14	1.7	1254	4	US-09-134-001C-1114	Sequence 1314, Ap
C 836	14	1.7	1024	4	US-09-328-475C-64	Sequence 64, Appl	C 909	14	1.7	1254	4	US-09-252-991A-7186	Sequence 7186, Ap
C 837	14	1.7	1024	4	US-09-328-475C-65	Sequence 65, Appl	C 910	14	1.7	1257	3	US-08-944-054-1	Sequence 1, Appl1
C 838	14	1.7	1024	4	US-09-328-475C-66	Sequence 66, Appl	C 911	14	1.7	1257	3	US-09-347-333-1	Sequence 1, Appl1
C 839	14	1.7	1026	4	US-09-328-352-3204	Sequence 3204, Ap	C 912	14	1.7	1260	4	US-09-166-350-22	Sequence 22, Ap
C 840	14	1.7	1032	4	US-09-134-001C-2647	Sequence 2647, Ap	C 913	14	1.7	1260	4	US-09-252-991A-1428	Sequence 1428, Ap
C 841	14	1.7	1032	4	US-09-107-532A-3057	Sequence 3057, Ap	C 914	14	1.7	1265	4	US-09-546-250-2	Sequence 2, Appl1
C 842	14	1.7	1034	3	US-09-363-970-35	Sequence 35, Appl	C 915	14	1.7	1279	4	US-09-185-244-2	Sequence 2, Appl1
C 843	14	1.7	1044	3	US-09-221-017B-555	Sequence 55, App	C 916	14	1.7	1279	4	US-09-471-913-6	Sequence 6, Appl1
C 844	14	1.7	1045	3	US-09-313-300-9	Sequence 9, Appl1	C 917	14	1.7	1287	3	US-08-945-258-11	Sequence 11, Appl
C 845	14	1.7	1053	2	US-08-927-307-1	Sequence 1, Appl1	C 918	14	1.7	1287	3	US-08-990-571-11	Sequence 11, Appl
C 846	14	1.7	1053	3	US-09-385-947-1	Sequence 1, Appl1	C 919	14	1.7	1287	4	US-08-723-142K-11	Sequence 11, Appl
C 847	14	1.7	1054	4	US-09-221-017B-609	Sequence 609, App	C 920	14	1.7	1287	4	US-09-528-784A-11	Sequence 11, Appl
C 848	14	1.7	1059	4	US-08-724-984A-3	Sequence 3, Appl1	C 921	14	1.7	1287	4	US-09-569-098A-11	Sequence 11, Appl
C 849	14	1.7	1059	4	US-09-517-605-8	Sequence 8, Appl1	C 922	14	1.7	1289	3	US-08-793-035-3	Sequence 3, Appl1
C 850	14	1.7	1070	2	US-08-179-557-17	Sequence 17, Appl	C 923	14	1.7	1299	3	US-09-546-250-3	Sequence 3, Appl1
C 851	14	1.7	1071	3	US-09-087-232A-14	Sequence 14, Appl	C 924	14	1.7	1303	4	US-08-894-440-2	Sequence 2, Appl1
C 852	14	1.7	1071	4	US-09-252-991A-7312	Sequence 7312, Ap	C 925	14	1.7	1303	4	US-09-458-093-2	Sequence 2, Appl1
C 853	14	1.7	1072	1	US-07-945-288-5	Sequence 5, Appl1	C 926	14	1.7	1307	3	US-09-282-305-17	Sequence 17, Appl
C 854	14	1.7	1072	1	US-08-462-831-5	Sequence 5, Appl1	C 927	14	1.7	1307	4	US-09-883-720-17	Sequence 17, Appl
C 855	14	1.7	1072	1	US-08-461-809-5	Sequence 5, Appl1	C 928	14	1.7	1309	4	US-09-544-618-5	Sequence 5, Appl1
C 856	14	1.7	1072	1	US-08-461-441-5	Sequence 5, Appl1	C 929	14	1.7	1314	4	US-09-134-001C-581	Sequence 581, App
C 857	14	1.7	1072	2	US-08-482-142-5	Sequence 5, Appl1	C 930	14	1.7	1314	4	US-09-328-352-4111	Sequence 4111, Ap
C 858	14	1.7	1072	2	US-08-478-572-5	Sequence 5, Appl1	C 931	14	1.7	1320	2	US-08-466-103A-1	Sequence 1, Appl1
C 859	14	1.7	1072	5	PCT-US93-08518-5	Sequence 5, Appl1	C 932	14	1.7	1333	4	US-09-325-932A-90	Sequence 90, Appl
C 860	14	1.7	1072	5	PCT-US93-08518-5	Sequence 5, Appl1	C 933	14	1.7	1333	4	US-09-328-352-3177	Sequence 3177, Ap
C 861	14	1.7	1083	4	US-09-107-532A-869	Sequence 869, App	C 934	14	1.7	1339	3	US-08-945-515-3	Sequence 3, Appl1
C 862	14	1.7	1086	4	US-09-134-001C-1208	Sequence 1208, Ap	C 935	14	1.7	1332	4	US-09-328-352-3988	Sequence 3988, Ap
C 863	14	1.7	1098	4	US-09-071-035-391	Sequence 35, App	C 936	14	1.7	1337	3	US-08-684-862-9	Sequence 9, Appl1
C 864	14	1.7	1098	4	US-09-252-991A-15756	Sequence 15756, A	C 937	14	1.7	1337	3	US-08-467-023-1	Sequence 1, Appl1
C 865	14	1.7	1101	4	US-09-134-001C-361	Sequence 361, App	C 938	14	1.7	1344	3	US-09-087-232A-16	Sequence 16, Appl
C 866	14	1.7	1103	1	US-08-553-633A-4	Sequence 4, Appl1	C 939	14	1.7	1359	4	US-08-591-468-21	Sequence 21, Appl
C 867	14	1.7	1104	4	US-09-328-352-3397	Sequence 3397, Ap	C 940	14	1.7	1359	5	PCT-US94-06430-21	Sequence 21, Appl
C 868	14	1.7	1108	2	US-08-852-807-2	Sequence 2, Appl1	C 941	14	1.7	1366	3	US-08-670-354-5	Sequence 5, Appl1
C 869	14	1.7	1110	4	US-09-252-991A-7127	Sequence 7127, Ap	C 942	14	1.7	1366	3	US-09-320-424-5	Sequence 5, Appl1
C 870	14	1.7	1110	4	US-09-328-352-3333	Sequence 3333, Ap	C 943	14	1.7	1366	4	US-09-825-563-5	Sequence 5, Appl1
C 871	14	1.7	1128	4	US-09-627-650B-12	Sequence 12, Appl	C 944	14	1.7	1366	5	PCT-US96-10895-5	Sequence 5, Appl1
C 872	14	1.7	1128	4	US-09-436-063C-12	Sequence 12, Appl	C 945	14	1.7	1371	1	US-09-620-312D-694	Sequence 694, App
C 873	14	1.7	1128	4	US-09-601-198-107	Sequence 107, App	C 946	14	1.7	1376	3	US-09-087-232A-12	Sequence 12, Appl
C 874	14	1.7	1131	4	US-09-328-352-2565	Sequence 2565, Ap	C 947	14	1.7	1376	4	US-09-016-434-1104	Sequence 1104, Ap
C 875	14	1.7	1134	4	US-09-502-528-2	Sequence 2, Appl1	C 948	14	1.7	1376	4	US-09-796-202-2	Sequence 2, Appl1
C 876	14	1.7	1140	4	US-09-227-357-53	Sequence 53, Appl	C 949	14	1.7	1380	4	US-09-328-352-3969	Sequence 3969, Ap
C 877	14	1.7	1143	4	US-09-328-352-1514	Sequence 1514, Ap	C 950	14	1.7	1383	3	US-08-935-263-3	Sequence 3, Appl1
C 878	14	1.7	1143	4	US-09-328-352-2115	Sequence 2115, Ap	C 951	14	1.7	1383	4	US-09-594-185-3	Sequence 3, Appl1
C 879	14	1.7	1146	4	US-09-134-001C-1921	Sequence 1921, Ap	C 952	14	1.7	1384	3	US-08-729-594A-33	Sequence 33, Appl
C 880	14	1.7	1146	4	US-09-328-352-8211	Sequence 2211, Ap	C 953	14	1.7	1384	4	US-08-937-993-133	Sequence 33, Appl
C 881	14	1.7	1151	3	US-09-475-316A-63	Sequence 63, Appl	C 954	14	1.7	1389	4	US-09-328-352-1154	Sequence 1154, Ap
C 882	14	1.7	1152	4	US-09-328-352-1843	Sequence 1843, Ap	C 955	14	1.7	1395	3	US-09-586-719-3	Sequence 3, Appl1
C 883	14	1.7	1158	4	US-09-328-352-1515	Sequence 1515, Ap	C 956	14	1.7	1395	4	US-09-449-476-88	Sequence 88, Appl
C 884	14	1.7	1164	4	US-09-134-001C-788	Sequence 788, App	C 957	14	1.7	1399	4	US-09-049-698-17	Sequence 17, Appl
C 885	14	1.7	1185	4	US-09-149-476-290	Sequence 290, App	C 958	14	1.7	1404	1	US-08-204-656B-3	Sequence 3, Appl1
C 886	14	1.7	1188	4	US-09-134-001C-623	Sequence 623, App	C 959	14	1.7	1404	1	US-08-204-656B-5	Sequence 5, Appl1
C 887	14	1.7	1200	4	US-09-071-035-389	Sequence 389, App	C 960	14	1.7	1404	1	US-08-467-831-4	Sequence 4, Appl1
C 888	14	1.7	1202	4	US-09-620-312D-97	Sequence 97, Appl	C 961	14	1.7	1404	1	US-08-470-702-2	Sequence 2, Appl1
C 889	14	1.7	1209	4	US-09-345-882-04	Sequence 24, Appl	C 962	14	1.7	1404	1	US-08-470-702-2	Sequence 2, Appl1
C 890	14	1.7	1215	4	US-09-328-352-49	Sequence 49, Appl	C 963	14	1.7	1404	1	US-08-470-702-4	Sequence 4, Appl1
C 891	14	1.7	1225	4	US-09-569-098A-99	Sequence 99, Appl	C 964	14	1.7	1404	1	US-08-467-831-2	Sequence 2, Appl1
C 892	14	1.7	1229	1	US-09-107-532A-2542	Sequence 2542, Ap	C 965	14	1.7	1404	1	US-08-467-831-3	Sequence 3, Appl1
C 893	14	1.7	1233	4	US-09-285-657-1	Sequence 1, Appl1	C 966	14	1.7	1404	1	US-08-467-831-4	Sequence 4, Appl1
C 894	14	1.7	1233	4	US-09-620-405B-492	Sequence 492, App	C 967	14	1.7	1410	4	US-09-328-352-124	Sequence 124, App
C 895	14	1.7	1236	2	US-08-379-556A-7	Sequence 7, Appl1	C 968	14	1.7	1414	3	US-08-466-343B-1	Sequence 1, Appl1
C 896	14	1.7	1236	2	US-08-741-134-5	Sequence 5, Appl1	C 969	14	1.7	1414	4	US-09-502-783A-1	Sequence 1, Appl1
C 897	14	1.7	1246	1	US-08-097-828-2	Sequence 2, Appl1	C 970	14	1.7	1425	2	US-09-328-352-586	Sequence 586, App
C 898	14	1.7	1246	2	US-08-480-756-2	Sequence 2, Appl1	C 971	14	1.7	1425	2	US-08-883-515-1	Sequence 1, Appl1
C 899	14	1.7	1246	2	US-08-462-403-2	Sequence 2, Appl1	C 972	14	1.7	1434	4	US-09-328-352-1302	Sequence 1302, Ap
C 900	14	1.7	1246	4	US-09-651-656-30	Sequence 30, Appl	C 973	14	1.7	1435	2	US-08-955-713-3	Sequence 3, Appl1
C 901	14	1.7	1246	4	US-09-650-855-30	Sequence 30, Appl	C 974	14	1.7	1437	4	US-09-107-532A-1268	Sequence 1268, Ap
C 902	14	1.7	1246	5	PCT-US93-10419-2	Sequence 2, Appl1	C 975	14	1.7	1442	4	US-08-833-752-3	Sequence 3, Appl1
C 903	14	1.7	1248	3	US-08-910-505-3	Sequence 3, Appl1	C 976	14	1.7	1443	4	US-09-252-991A-12097	Sequence 12097, A

977 14 1.7 1446 4 US-09-593-359-1 Sequence 1, Appl1  
C 978 14 1.7 1458 4 US-09-134-001C-989 Sequence 999, App  
C 979 14 1.7 1477 4 US-08-833-752-2 Sequence 2, Appl1  
980 14 1.7 1479 4 US-09-107-532A-3408 Sequence 3408, Ap  
C 981 14 1.7 1479 4 US-09-107-532A-3408 Sequence 3408, Ap  
C 982 14 1.7 1483 5 PCT-US93-06251-88 Sequence 88, Appl  
C 983 14 1.7 1484 5 US-08-185-828A-22 Sequence 22, Appl  
984 14 1.7 1485 4 US-08-936-165A-218 Sequence 218, Appl  
985 14 1.7 1488 4 US-09-328-352-1977 Sequence 3977, Ap  
986 14 1.7 1492 4 US-09-484-970B-140 Sequence 140, App  
C 987 14 1.7 1494 4 US-09-634-238-7 Sequence 7, Appl1  
988 14 1.7 1497 4 US-09-134-001C-1824 Sequence 1824, Ap  
C 989 14 1.7 1500 4 US-09-252-991A-11439 Sequence 11439, A  
C 990 14 1.7 1508 3 US-08-660-347-1 Sequence 1, Appl1  
C 991 14 1.7 1522 4 US-09-413-574-1 Sequence 1, Appl1  
992 14 1.7 1522 4 US-09-620-312D-296 Sequence 296, App  
C 993 14 1.7 1526 4 US-09-482-273-33 Sequence 33, Appl  
C 994 14 1.7 1532 2 US-08-185-828A-11 Sequence 11, Appl  
C 995 14 1.7 1541 4 US-09-016-434-775 Sequence 775, App  
996 14 1.7 1542 4 US-09-134-001C-1653 Sequence 1653, Ap  
997 14 1.7 1543 3 US-08-714-918-82 Sequence 82, Appl  
998 14 1.7 1543 3 US-09-265-315-82 Sequence 82, Appl  
999 14 1.7 1543 3 US-09-265-315-82 Sequence 82, Appl  
1000 14 1.7 1543 3 US-09-266-417-82 Sequence 82, Appl

## ALIGNMENTS

RESULT 1  
US-09-557-884-1/c  
; Sequence 1, Application US/09557884  
; Patent No. 6506581  
; GENERAL INFORMATION:  
; APPLICANT: Fleischmann et al.  
; TITLE OF INVENTION: The Nucleotide sequence of  
; the Haemophilus influenzae Rd Genome, Fragments  
; thereof, and Uses Thereof  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: MD  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3 1/2 inch diskette  
; COMPUTER: Dell Pentium  
; OPERATING SYSTEM: MS DOS v6.22  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/557, 884  
; FILING DATE: 25-Apr-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/476,102  
; FILING DATE: JUN-5-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Michelle S. Marks  
; REGISTRATION NUMBER: 41, 971  
; REFERENCE/DOCKET NUMBER: PB186P3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-309-8504  
; TELEFAX: 301-309-8439  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1830121 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-557-884-1

Query Match 2.5%; Score 21; DB 4; Length 1830121;  
Best Local Similarity 100.0%; Pred. No. 0.54;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 788 TTTCGCTGAAATGACTTAT 808  
Db 1174328 TTTCGCTGAAATGACTTAT 1174308

## RESULT 2

US-09-643-990A-1/c  
; Sequence 1, Application US/09643990A  
; Patent No. 6528289  
; GENERAL INFORMATION:  
; APPLICANT: Robert D. Fleischmann  
; Mark D. Adams  
; Owen White  
; Hamilton O. Smith  
; J. Craig Venter  
; TITLE OF INVENTION: The Nucleotide sequence of  
; the Haemophilus influenzae Rd Genome, Fragments  
; thereof, and Uses Thereof  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville,  
; STATE: MD  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3 1/2 inch diskette  
; COMPUTER: Dell Pentium  
; OPERATING SYSTEM: MS DOS v6.22  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/643, 990A  
; FILING DATE: 23-Aug-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/487, 429  
; FILING DATE: 1995-06-07  
; APPLICATION NUMBER: 08/426, 787  
; FILING DATE: 1995-04-21  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kenley K. Hoover  
; REGISTRATION NUMBER: 40, 302  
; REFERENCE/DOCKET NUMBER: PB186P1C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-610-5790  
; TELEFAX: 310-309-8439  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1830121 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-643-990A-1  
Query Match 2.5%; Score 21; DB 4; Length 1830121;  
Best Local Similarity 100.0%; Pred. No. 0.54;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 788 TTTCGCTGAAATGACTTAT 808  
Db 1174328 TTTCGCTGAAATGACTTAT 1174308  
RESULT 3  
US-09-453-702B-30  
; Sequence 30, Application US/09453702B

Patent No. 6365723  
GENERAL INFORMATION:  
APPLICANT: Blatner, Frederick R.  
Burland, Valerie  
Perna, Nicole T.  
Plunkett, Guy  
Welch, Rod  
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Quarles & Brady  
STREET: 1 South Pinckney Street  
CITY: Madison  
STATE: WI  
COUNTRY: US  
ZIP: 53701-2113  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 8.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/453,702B  
FILING DATE: 03-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/110,955  
FILING DATE: 04-DEC-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Seay, Nicholas J.  
REGISTRATION NUMBER: 27386  
REFERENCE/DOCKET NUMBER: 960296.95017  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 251-5000  
TELEFAX: (608) 251-9166  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1394  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 30:  
US-09-453-702B-30  
Query Match 2.3%; Score 19; DB 4; Length 1394;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 40 GGTTGATGAGCTGATGAAG 58  
Db 612 GGTTGATGAGCTGATGAAG 630  
RESULT 4  
US-09-212-247C-1/c  
Sequence 1, Application US/09212247C  
Patent No. 6391603  
GENERAL INFORMATION:  
APPLICANT: POMPEJUS, Markus; SUELBINGER, Harald; JOEFFKEN, Hans  
Wolfgang; DOVAL, Jose Luis Revuelta; JIMENEZ, Alberto;  
and GARCIA, Maria Angeles Santos  
TITLE OF INVENTION: Genes of purine biosynthesis from Ashbya Gossypii  
and the use thereof in microbial riboflavin  
synthesis  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kell & Weinkauff  
STREET: 1101 Connecticut Avenue  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: IBM AT-compatible, Pentium processor  
OPERATING SYSTEM: Windows 95  
SOFTWARE: WordPerfect version 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/212,247C  
FILING DATE: 16-Dec-1998  
CLASSIFICATION: <Unknown>  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1911 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..625  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 626..1582  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 1583..1911  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-212-247C-1  
Query Match 2.3%; Score 19; DB 4; Length 1911;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 767 AAGAGTACAAATGATGA 785  
Db 247 AAGAGTACAAATGATGA 229  
RESULT 5  
US-09-699-266A-5  
Sequence 5, Application US/09699266A  
Patent No. 6559354  
GENERAL INFORMATION:  
APPLICANT: Ma, Hongchang  
APPLICANT: Morakinyo, Layo O.  
APPLICANT: Odell, Joan T.  
APPLICANT: Orozco Jr., Emil M.  
APPLICANT: Rafalski, J. Antoni  
TITLE OF INVENTION: TRANSCRIPTION AND GENE EXPRESSION REGULATORS  
FILE REFERENCE: B1164 US NA  
CURRENT APPLICATION NUMBER: US/09/699,266A  
CURRENT FILING DATE: 2000-10-27  
PRIOR APPLICATION NUMBER: PCT/US99/08385  
PRIOR FILING DATE: 1999-04-16  
PRIOR APPLICATION NUMBER: 60/083,212  
PRIOR FILING DATE: 1998-04-27  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 5  
LENGTH: 999  
TYPE: DNA  
ORGANISM: Triticum aestivum  
US-09-699-266A-5  
Query Match 2.2%; Score 18; DB 4; Length 999;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 598 ATTGAAGAGATGAAGAG 615  
Db 836 ATTGAAGAGATGAAGAG 853

```
RESULT 6
US-09-610-185C-3
; Sequence 3, Application US/09610185C
; Patent No. 6573432
; GENERAL INFORMATION:
; APPLICANT: Borevitz, Justin
; APPLICANT: Xia, Yijl
; APPLICANT: Dixon, Richard A.
; APPLICANT: Lamb, Christopher J.
; TITLE OF INVENTION: REGULATION OF ANTHOCYANIN PIGMENT
; FILE REFERENCE: SALKINS.003C1
; CURRENT APPLICATION NUMBER: US/09/610.185C
; CURRENT FILING DATE: 2000-07-05
; PRIOR APPLICATION NUMBER: US 09/603,244
; PRIOR FILING DATE: 2000-06-23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1033
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-610-185C-3

Query Match      2.2%; Score 18; DB 4; Length 1033;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      78 ATGATTGATAGTATG 95
Db      181 ATGATTGATAGTATG 198

RESULT 7
US-08-928-941D-28/c
; Sequence 28, Application US/08928941D
; Patent No. 6180763
; GENERAL INFORMATION:
; APPLICANT: Hirai, Hiroshi
; APPLICANT: Sherr, Charles
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928.941D
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-002 N CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3767 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-928-941D-28

Query Match      2.2%; Score 18; DB 3; Length 3767;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      609 TGAAGAGTTGTATTAA 626
Db      2620 TGAAGAGTTGTATTAA 2603

RESULT 8
US-08-928-941D-30/c
; Sequence 30, Application US/08928941D
; Patent No. 6180763
; GENERAL INFORMATION:
; APPLICANT: Hirai, Hiroshi
; APPLICANT: Sherr, Charles
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928.941D
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-002 N CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3767 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-928-941D-30

Query Match      2.2%; Score 18; DB 3; Length 3767;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      609 TGAAGAGTTGTATTAA 626
Db      2620 TGAAGAGTTGTATTAA 2603

RESULT 9
US-09-280-590A-28/c
```

; Sequence 28, Application US/09280590A  
; Patent No. 6303772  
; GENERAL INFORMATION:  
; APPLICANT: Hirai, Hiroshi  
; Sherr, Charles  
; Inoue, Kazushi  
; Bodner, Sarah M.  
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES  
; THEREOF  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/280,590A  
; FILING DATE: 29-Mar-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1340-1-002 N CP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3767 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
; US-09-280-590A-28  
; Query Match 2.2%; Score 18; DB 4; Length 3767;  
; Best Local Similarity 100.0%; Pred. No. 21;  
; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
; QY 609 TGAAGAGTTGTAATTAA 626  
; DB 2620 TGAAGAGTTGTAATTAA 2603  
; RESULT 10  
; US-09-280-590A-30/c  
; Sequence 30, Application US/09280590A  
; Patent No. 6303772  
; GENERAL INFORMATION:  
; APPLICANT: Hirai, Hiroshi  
; Sherr, Charles  
; Inoue, Kazushi  
; Bodner, Sarah M.  
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES  
; THEREOF  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; Floor  
; CITY: Hackensack

; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/280,590A  
; FILING DATE: 29-Mar-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1340-1-002 N CP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3767 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: RNA  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; SEQUENCE DESCRIPTION: SEQ ID NO: 30:  
; US-09-280-590A-30

; Query Match 2.2%; Score 18; DB 4; Length 3767;  
; Best Local Similarity 100.0%; Pred. No. 21;  
; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
; QY 609 TGAAGAGTTGTAATTAA 626  
; DB 2620 TGAAGAGTTGTAATTAA 2603

; RESULT 11  
; US-09-328-352-1695  
; Sequence 1695, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 1695  
; LENGTH: 7515  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
; US-09-328-352-1695

; Query Match 2.2%; Score 18; DB 4; Length 7515;  
; Best Local Similarity 100.0%; Pred. No. 21;  
; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; QY 340 AAGTTAATCTACTATAA 357  
; DB 5008 AAGTTAATCTACTATAA 5025

; RESULT 12  
; US-08-926-842B-11/c  
; Sequence 11, Application US/08926842B  
; Patent No. 6030807  
; GENERAL INFORMATION:

APPLICANT: Sa-No. 6030807neira, Isabel  
APPLICANT: de Lencastre, Herminia  
TITLE OF INVENTION: HIGHLY REGULABLE PROMOTER FOR HETEROLOGOUS GENE  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 64  
CORRESPONDENCE ADDRESSES: 64  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/926,842B  
FILING DATE: 10-SEP-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-089 N  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10917 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Bacillus subtilis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 228..1718  
OTHER INFORMATION: /product= "araA"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1732..3417  
OTHER INFORMATION: /product= "araB"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3431..4120  
OTHER INFORMATION: /product= "araD"  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 4107..4916  
OTHER INFORMATION: /product= "araL"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 4913..6097  
OTHER INFORMATION: /product= "araW"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 6128..7429  
OTHER INFORMATION: /product= "araN"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 7465..8406  
OTHER INFORMATION: /product= "araP"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 8407..9255  
OTHER INFORMATION: /product= "araQ"  
FEATURE:  
NAME/KEY: CDS

LOCATION: 9271..10773  
OTHER INFORMATION: /product= "abfA"  
US-08-926-842B-11  
Query Match 2.2%; Score 18; DB 3; Length 10917;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 781 GATGATTTTCGCTGAA 798  
DB 6210 GATGATTTTCGCTGAA 6193  
RESULT 13  
US-09-313-294A-3585/C  
Sequence 3585, Application US/09313294A  
Patent No. 6476212  
GENERAL INFORMATION:  
APPLICANT: Lalagudi, Raghunath V.  
APPLICANT: Ico, Laura Y.  
APPLICANT: Sherman, Bradley K.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
FILE REFERENCE: PL-0017 US  
CURRENT APPLICATION NUMBER: US/09/313,294A  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 7600  
SOFTWARE: PERL Program  
SEQ ID NO 3585  
LENGTH: 271  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6476212 700611935H1  
NAME/KEY: unsure  
LOCATION: 166-184  
OTHER INFORMATION: a, t, c, g, or other  
US-09-313-294A-3585  
Query Match 2.1%; Score 17; DB 4; Length 271;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 100 GGAATGCGATCTGT 116  
DB 155 GGAATGCGATCTGT 139  
RESULT 14  
US-09-328-352-1167  
Sequence 1167, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 1167  
LENGTH: 750  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-1167  
Query Match 2.1%; Score 17; DB 4; Length 750;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 206 TTGAACAAGATGAAGTG 222  
DB 248 TTGAACAAGATGAAGTG 264



## RESULT 15

US-08-753-007A-7  
; Sequence 7, Application US/08753007A  
; Patent No. 6074841  
; GENERAL INFORMATION:  
; APPLICANT: Gearing, David P.  
; APPLICANT: Busfield, Samantha J.  
; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES  
; TITLE OF INVENTION: AND USES THEREFOR  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/753,007A  
; FILING DATE: 19-NOV-1996  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/699,591  
; FILING DATE: 19-AUG-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Faase, J. Peter  
; REGISTRATION NUMBER: 32,983  
; REFERENCE/DOCKET NUMBER: 07334/022001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1476 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 69...1475  
; OTHER INFORMATION:  
; US-08-753-007A-7

Query Match 2.1%; Score 17; DB 3; Length 1476;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 599 TTGAAGAAGATGAAGAG 615  
|||||  
Db 174 TTGAAGAAGATGAAGAG 190  
|||||

Search completed: January 30, 2004, 10:47:57  
Job time : 94 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 05:45:21 ; Search time 2209 Seconds

(without alignments)  
15278.607 Million cell updates/sec

Title: US-10-033-190-1

Perfect score: 825

Sequence: 1 atgaacagtcacatcatgtc.....tatgaatcacttgatcaa 825

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapept 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_cm:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_dr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rtd:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	264.8	32.1	1034	8 AF146702	AF146702 Petunia x
2	261.6	31.7	866	8 AF146704	AF146704 Petunia x
3	260.8	31.6	868	8 AF146706	AF146706 Petunia x
4	260.8	31.6	868	8 AF146707	AF146707 Petunia x
5	260	31.5	865	8 AF146703	AF146703 Petunia x
6	251.4	30.5	909	8 AF146705	AF146705 Petunia x
7	247.8	30.0	865	8 AF146709	AF146709 Petunia x
8	244.4	29.6	858	8 AF146708	AF146708 Petunia x
9	200.4	24.3	920	8 AB073010	AB073010 Vitis lab
10	200.4	24.3	1173	8 AB073013	AB073013 Vitis lab
11	195.6	23.7	903	8 AB073012	AB073012 Vitis lab
12	189.6	23.0	420	8 AY008379	AY008379 Arabidops
13	186	22.5	992	8 AF325124	AF325124 Arabidops
14	186	22.5	1033	6 AF358283	AF358283 Sequence
15	186	22.5	1033	8 AF062915	AF062915 Arabidops
16	185.4	22.5	741	8 AY008378	AY008378 Arabidops
17	180.2	21.8	956	6 AX358281	AX358281 Sequence
18	179	21.7	747	6 AX507939	AX507939 Sequence
19	179	21.7	747	6 AX651641	AX651641 Sequence
20	179	21.7	925	8 AF062908	AF062908 Arabidops
21	165.2	20.0	1133	8 AF358284	AF358284 Sequence
22	154.8	18.8	774	6 AX507941	AX507941 Sequence
23	154.8	18.8	774	8 BT000841	BT000841 Arabidops
24	154.8	18.8	897	8 AF062859	AF062859 Arabidops
25	154.8	18.8	927	8 AY072543	AY072543 Arabidops
26	151.8	18.4	1135	8 AF401220	AF401220 Fragaria
27	150	18.2	1462	8 PM039448	PM039448 Picea maria
28	147.2	17.8	687	8 AF495524	AF495524 Arabidops
29	147	17.8	742	8 BT0005085	BT0005085 Arabidops
30	147	17.8	764	8 ATT026936	ATT026936 Arabidops
31	147	17.8	871	8 AY086813	AY086813 Arabidops
32	147	17.8	933	8 BT004171	BT004171 Arabidops
33	146.8	17.8	639	8 AF207991	AF207991 Arabidops
34	146.6	17.8	991	8 AF336285	AF336285 Gossypium
35	144.4	17.5	1151	8 AF034134	AF034134 Gossypium
36	143.8	17.4	1006	8 COMMYBA	L04497 Gossypium h
37	143.8	17.4	1168	8 AF336282	AF336282 Gossypium
38	143	17.3	810	6 AX507651	AX507651 Sequence
39	143	17.3	810	6 AX651705	AX651705 Sequence
40	143	17.3	844	8 ATT026937	ATT026937 Arabidops
41	142.4	17.3	884	8 AF034130	AF034130 Gossypium
42	141.8	17.2	579	8 AF502295	AF502295 Cucumis s
43	140.2	17.0	967	8 ATWYBRFP	Z68158 A. thaliana
44	139.6	16.9	1331	8 AY087252	AY087252 Arabidops
45	138	16.7	960	6 AX507443	AX507443 Sequence

#### ALIGNMENTS

RESULT 1

AF146702

LOCUS AF146702 1034 bp mRNA linear PLN 01-MAY-2000

DEFINITION Petunia x hybrida An2 protein (an2) mRNA, an2-V26 allele, complete cds.

ACCESSION AF146702

VERSION AF146702.1 GI:7673083

KEYWORDS

SOURCE

ORGANISM

Petunia x hybrida

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamiales; Solanales; Solanaceae; Petunia.

REFERENCE

1 (bases 1 to 1034)

Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N.,

Pred. No. is the number of results predicted by chance to have a

TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
Mol, J. and Koes, R.					Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color			
	Plant Cell 11 (8), 1433-1444 (1999)							
	99380006							
	10449578							
	2 (bases 1 to 1034)							
	Quattrocchio, F., Wang, J., van der Woude, K., Souer, E., de Vetten, N., Mol, J. and Koes, R.							
	Direct Submission							
	Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan 1087, Amsterdam 1081HV, Netherlands							
	Location/Qualifiers							
	1. .1034							

```
gene 1. .1034
      /gene="an2"
      /allele="an2-V26"
CDS 100. .867
     /start="1"
     /end="867"
```

BASE COUNT	365 a	154 c	216 g	299 t
ORIGIN				

Query Match	32.1%;	Score 264.6;	DB 8;	Length 1034;
Best Local Similarity	67.8%;	Pred. No. 3.5e-42;		
Matches 431; Conservative	0;	Mismatches 172;	Indels 33;	Gaps 3;

Qy	1	ATGAAACGATCATCATGTCTTCAATTGGAGATGAGAAAAGGTTACATGACATGAAACAA	60
Db	100	ATGAGTACTTTAAATGCATCAACATCAGAGATGAGAAAAGGTGCATGACCGAGAGACA	159
Qy	61	GATTTCCTTCTAAAGAAAATGTATTGATAGTATGGTGAAGAAAATGCACTCTTGTCCC	120
Db	160	GATCTTTTATTGAGAGAAATGCATTGACAAAGTATGAGAGAAAGATGGCATCTAGTTCCA	219
Qy	121	ATAAGACTGTCTGAAATGATAGTGTGCGAAAAGTGTAGATTGAGGTGCTGAATTATCTA	180
Db	220	GTTAGAGCTGTCTGAAATGATAGTGCAGGAAAAGTTGCACACTTGAAGTGTGAATTATCTA	279
Qy	181	AGGCCACATATCAAGAGAGTGACTTTGAAACAAGTGAAGTGATCTCATTTTAGGCTT	240
Db	280	AGGCCACATATPAAAAAGGGGACCTTCTTTGATGAAGTAGAGATCTCATTTTAGGCTT	339
Qy	241	CATAGACTTTAGGCAAGATGGTCACTTATTTGTGTAGACTCCCGAAGGACAGCT	300
Db	340	CATAGACTTTAGGCAAGATGGTCACTTATAGCTGTAGACTTCTGGAAGAACACCA	399
Qy	301	AACGATGTGAAAACTATTGGACAACACTATCTTTAAAGAAATTAATACTACTAAAAAT	360
Db	400	AACGATGTCAAAAACTATTGGAAACCCACCTTGCAAAAGATTAATGTCTCTCA-----	455
Qy	361	GTTCCTGCGAAAAAGATTAAACATATAGTGTGAGAAATTAAGTACTTAAGATTGAATTATA	420
Db	456	--TGATCGAAACAAGAGACAGACAAGAACCGTGAATAATTAACGAGAACACATATATA	513
Qy	421	AAACTCAGACGCAAGTATTTCTCAAGCACATGAAGAATGTTCAAA-----	470

Db	514	AAACCTCGCTCGGACCTTCTCAAGGCGCGCANTGAATAATTTCTTGTGAAAGCG	573
Qy	471	-----CAATATGTAAATTTTGGACGAGAGGAACTTGCACAGAAATATATA-----	517
Db	574	AAAAGTTGTAAATATAAACAATAATGACAGAAATAGAGTGACACAGAAATATAAAGTTT	633
Qy	518	----GTGATAAACAAATCCAGATGCATGTCAGCAACGTAGATCATGTGTGATTAAT	573
Db	634	AGTGAATGAAAGCAAAAAACCGAGAAATCGATATATATGACTTCAATGCTGGCCAAAT	693
Qy	574	TTACTGAAATATTGCATATGACGATTTATGAAGAAGAT	609
Db	694	TTTATTAGCCCAACAACATTGAGATTGAGAAATTAGTT	729

RESULT 2	AF146704	866 bp	mRNA	linear	PLN 01-MAY-2000
LOCUS	AF146704				
DEFINITION	AF146704 Petunia integrifolia An2 protein (an2) mRNA, an2-86 allele, complete cds.				
ACCESSION	AF146704				
VERSION	AF146704.1	GI:7673087			
KEYWORDS					
SOURCE	Petunia integrifolia				
ORGANISM	Petunia integrifolia				

REFERENCE  
1 (baees 1 to 866)

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 866)	Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N.,...	Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color	Plant Cell 11 (8), 1433-1444 (1999)

REFERENCE 2 (bases 1 to 866)  
AUTHORS Quattrocchio, F., Wang, J., van der Woude, K., Souer, E., de Vetten, N.,  
Mol, J., and Koeg, R.  
TITLE Direct Submission  
JOURNAL Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan  
1067, Amsterdam 1081HV, Netherlands

**Source**

```

/organism="Petunia integrifolia"
/mol_type="mRNA"
/cultivar="S6"
/db_xref="taxon:4103"
/chromosome="6"
/rnause_type="petal_limb"
1..866..
/gene="an2"
/allele="an2-S6"
12..779..
-CDS

```

BASE COUNT  
ORIGIN

Query Match	31.7%;	Score 261.6;	DB 8;	Length 866;
Best Local Similarity	67.5%;	Pred. No. 1.5e-41;		
Matches 429;	Conservative 0;	Mismatches 174;	Indels 33;	Gaps 3

OY		1	ATGAACAGTAACTCATATCTCTTCATTGGGAGTGAGAATAAAGGTCTATGGACTGATGAAGAA	60
Db		12	ATGAGTACTTTCTTAATGCATCAACATCAGAGTAGAATAAAGGTGCATGAGCCGAGGAGAA	71
OY		61	GATTTTCTCTTAAGAAAATGTAATGATAGATGCGTGAAGAAAAATGCGATCTTGTGCC	120
Db		72	GATCTTTATTTAGAGAAATGCAATTCAGAAAGTATGGAAGAAAGAAATGGCATCTAGTTCCA	131
OY		121	ATTAAGACTGGTCTCGAATATGATGTCTGGAAAAAGTTTAAATTGAGGTGCTGAATTTCTA	180
Db		132	GTTAGAGCTGTCTGTAATAGATGCGAAAAAGTTGCACAATTAGGTGTGTAATTAATCTA	191
OY		181	AGGCCACATATCAAGAGAGGTGACTTTGAACAAGATGAAGTGAATCTGATTTTGAGGCTT	240
Db		192	AGGCCACATATTAATAAGAGGGGACTTCTCTTTGGATGAATGATCTCATTTTGAAGGCTT	251
OY		241	CATAGAGCTTTAGGCAACAGATGCTCATCTTATGCTGTAGACTTCCCGAAGACAGACT	300
Db		252	CATAGAGCTTTAGGCAACAGATGCTCATCTTATGCTGTAGACTTCCCGAAGAACAGCA	311
OY		301	AACGATGTGAAAACTATTGGAAACCTAATCTTTTAAGGAAGTTAAATACCTAATAAATT	360
Db		312	AACGATGTCAAAAACTATTGGAAACCCACACTTCGAAAGAAATTAATGTCTCTCA----	367
OY		361	GTTCTCTGCGAAAAAGTTAAACATAATAGTGTGGAAGAAATTAGTACTAABATTAATTTA	420
Db		368	--TGATCGAAACACAGAGACAGAAACAAAGCCATGAATAATTCGGAACACACATAATA	425
OY		421	AAACCTCACAGACCGAGATATTTCTCAAGCACAATGAGAAATGTT-----	465
Db		426	AAACCTCTCTCTCAGACCTTCTCAAGCGCGGCAATGAATCATGTTTTCTTTGGAAACGC	485
OY		466	---ACAAACAATATATGTAATTTTGGACGAGAGAGAACTTCGAAAGAAATATAA-----	517
Db		486	AAAGTTGTAGTAAAAACATATAGACAAGATGAAGGTGACACCGAATATATAAGTTT	545
OY		518	----GTGGAAGAAACAACTCCAGATGCAATGATGACAACTGATGATCCATGSGTGAATAAT	573
Db		546	AGTATGAGGAACAAACCGGAAAGATCGATGATGATGAGACTTCAATGTGTGGCAAAAT	605
OY		574	TTACTGGAATAATTCGATGACGATTTGAGAGAGAT	609
Db		606	TTATTTAGCAAACACATTTGAGATTTGAGAAATTAGTT	641
RESULT 3				
AF146706			868 bp mRNA linear PLN 01-MAY-2000	
LOCUS			Petunia x hybrida An2 truncated protein (an2) mRNA, an2-W22 allele,	
DEFINITION			complete cds.	
ACCESSION			AF146706	
VERSION			AF146706.1 GI:7673091	
KEYWORDS				
SOURCE				
ORGANISM			Petunia x hybrida	
			Petunia x hybrida	
			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
			Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;	
			Assteridae; lamids; Solanales; Solanaceae; Petunia.	
REFERENCE			1 (bases 1 to 868)	
AUTHORS			Quattrocchio,F., Wing,J., van der Woude,K., Souer,E., de Vetten,N.,J.	
TITLE			Molecular analysis of the anthocyanin2 gene of petunia and its role	
JOURNAL			Plant Cell 11 (8), 1433-1444 (1999)	
MEDLINE			99380006	
PUBMED			10449578	
REFERENCES			2 (bases 1 to 868)	
AUTHORS			Quattrocchio,F., Wing,J., van der Woude,K., Souer,E., de Vetten,N.,J.	
TITLE			Direct Submission	
JOURNAL			Submitted (29-Apr-1999) Genetics, Vrije Universiteit, De Boelelaan	
FEATURES			1087, Amsterdam 1081HV, Netherlands	
			Location/Qualifiers	

source	1. .868	/organism="Perunia x hybrida"
		/mol_type="mRNA"
		/cultivar="W22"
		/db_xref="taxon:4102"
		/chromosome="6"
		/tissue_type="petal limb"
gene	1. .868	/gene="an2"
		/allele="an2-W22"
CDS	12. .395	/gene="an2"
		/note="hanchocyanin regulator; myb-related transcription factor"
		/codon_start=1
		/product="An2 truncated protein"
		/protein_id="AAE6731.1"
		/db_xref="GI:7673092"
		/translation="MSTSNASTGVKRGAWTEERDILLRECIDYGESEKMLVPRARINRRKRSCLRRLMYLRPHIKRQDFSLDEVDLILRLKLLGNRSLIAGRLPGRTANIVKQVNTWTLRKRLIAPHDQKESK"
		392. .395
variation		/gene="an2"
		/note="causes frameshift leading to truncated protein"
		/replace="nnn"
BASE COUNT	310 a 140 c 191 g 227 t	
ORIGIN		
Query Match	31.6%; Score 260.8; DB 8; Length 868;	
Best Local Similarity	67.0%; Pred. No. 2.2e-41;	
Matches 429; Conservative	0; Mismatches 182; Indels 29; Gaps 3.	
Qy	1 ATGAACAGTACATGTATGCTTTCATTTGGAGATGAGAAAAGTTTCATGACTGATGAAGA 60	
Db	12 ATGAGTACTTCTATGATCAATCAATCAGAGTAAGAAAGTGCATGACCGAGAGAA 71	
Qy	61 GATTTCTCTTCAAGAAAATGATATGATATGATGATGAGAAAATGGCATCTTGTGCC 120	
Db	72 GATCTTTTATTTGAGAGATGCATTGACACAGATGAGAGAAAGTGCATCTTGTGTCCA 131	
Qy	121 ATAAAGCTGGTCTGAATAGATGTCCGAAAAGTTGTAGATTGAGGTGGCTGAATTATCTA 180	
Db	132 GTTAGAGCTGTCTGAATAGATGATGAGAGAAAAGTTGCACACTTAGGTGTTGATATCTTA 191	
Qy	181 AGGCACATATCAAGAGAGGTGACTTTTGAACAAATGAAGGATCTTCATTTTGGAGCTT 240	
Db	192 AGGCCACATATTAAGAGAGGAGGACTTCTCTTTGGATGAAGATCATCTATTTAGAGCTT 251	
Qy	241 CATAGCTCTTAGGCAACAGATGCTCACTTATTCGTGTAGACTTCCCGAAGACAGCT 300	
Db	252 CATAGGCTTCTAGGCAACAGATGCTCACTTATTCCTGGAACACTTCCCGAAGAACTGCA 311	
Qy	301 AACGATGGAACAACTATTTGGAACAACTAATCTTCTTAGAGAGTTAAATCACTAAATT 360	
Db	312 AACGATGCAAAACTATTTGGAACAACTGACCTTGGAAAAGTTAATGCTTCATCAT 371	
Qy	361 GTTCTCGCGAAAAGATTAACTAATAGTGTGAGAAAATTAGTACTAGATTGAATTATA 420	
Db	372 CAGAAACAAAGACAAAGAGCTACCAAAAGCAATGAA--ATCACGAGAACAACTATATA 429	
Qy	421 AAACCTCAACGACGAAGTATTTCTCAAGCAACTGAAGAATGTT----- 465	
Db	430 AAACCTCGTCTCGGACCTTCTCAAGCCGCGCAATGAAATCATGTTTCTTTGTGGAACGGC 489	
Qy	466 ---ACAAACAATATGTAATTTTGGAGAGAGAACTTTCAGAGAAATATAA----- 517	
Db	490 AAAAGTTGTAATATAAACACTATAGACAAAGATGAAGGTGACACAGAAATATATAAGTTT 549	
Qy	518 ----GTGAGAAAACAACTCGATGATGCAATGAGAACATGATGCAATGCTGTGATAAT 573	
Db	550 ACTGATGAGAGCAAAAACCGAAGATTTCATTAATGATGACCTTCATGATGTCGCCAAT 609	
Qy	574 TTACTGAAAATTGCAATGACGATATTGGAAGAGATGAAG 613	

Db	610	TTATTAGCCAAACATTCGACATTAGGACTTAACTAATG	649
RESULT 4	AF146707	668 bp	linear
LOCUS	AF146707		PLN 01-MAY-2000
DEFINITION	Petunia x hybrida An2 truncated protein (an2) mRNA, an2-W44 allele,		
ACCESSION	AF146707		
VERSION	AF146707.1		
KEYWORDS	GI:7673093		
SOURCE	Petunia x hybrida		
ORGANISM	Petunia x hybrida		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamiales; Solanales; Solanaceae; Petunia.		
AUTHORS	1 (bases 1 to 868) Quattrocchio, F., Ming, J., van der Woude, K., Souer, E., de Vetten, N., Moij, J. and Koes, R.		
TITLE	Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color		
JOURNAL	Plant Cell 11 (8), 1433-1444 (1999).		
MEDLINE	99380006		
PUBMED	10449578		
REFERENCES	2 (bases 1 to 868) Quattrocchio, F., Ming, J., van der Woude, K., Souer, E., de Vetten, N., Moij, J. and Koes, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-Apr-1999) Genetics, Vrije Universiteit, De Boelelaan 1087, Amsterdam 1081HV, Netherlands		
FEATURES	Location/Qualifiers		
source	1..868		
organism	"Petunia x hybrida"		
mol_type	"mRNA"		
culivar	"W44"		
db_xref	"taxon:4102"		
chromosome	"6"		
tissue_type	"petal 1mb"		
gene	1..868		
gene	/gene="an2"		
allele	"an2-W44"		
gene	12..395		
gene	/gene="an2"		
note	"anthocyanin regulator; myb-related transcription factor"		
codon_start	1		
product	"an2 truncated protein"		
protein_id	"AA66732.1"		
db_xref	"GI:7673094"		
translation	"MSTSNASTSGVRKAWTEEDLLRECIERYGSGKMLVPRVAG LNRKRSCLRLNLYLRPHIRKGFSLDEVELIRLHLKLRWSLLAGRLPRTAND VKNRWNTLRLKLLIAPHQKESKS"		
variation	392..395		
gene	/gene="an2"		
note	"causes frameshift leading to truncated protein"		
replace	"nnn"		
BASE COUNT	307 a 137 c 195 g 229 t		
ORIGIN			
Query Match	31.6%; Score 260.8; DB 8; Length 868;		
Best Local Similarity	67.0%; Pred. No. 2.2e-41;		
Matches 429; Conservative 0; Mismatches 182; Indels 29; Gaps 3;			
Db	1	ATGAACAGTACATCTATGTCCTTCATTGGAGTGAAGAAAGGTCATGACATGAAGAA	60
Db	12	ATGAGTACTCTTAATGATCAACATCAAGAGTGAAGAAAGTGATGACGAGAGAA	71
Qy	61	GATTTCTCTTAAGAAATGATTTAGTATGATGATGAAGAAATGATCTTTGTTCC	120
Db	72	GATCTTTATGAGAGATGATGATGAGATGAGAGAGAGATGATGATGATGATGATG	131
Qy	121	ATTAAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	180

Db		132	GTTAGAGCTGCTGTGAATATGACAGAAAAAGTTGCAGACTTAGTGCGTATTGAANTTAATCTA	191
OY		181	AGGCCACAATTCAGAGAGSTGACTTTGAACAAGATGAAGTGGATCTCATTTTGAGGCTT	240
Db		192	AGGCCACATATAAAAAGAAGGGGACTTCTCTTTGGATGMACTAAGCTTATTTTGAGGCTT	251
OY		241	CATAAGCTCTTAGSCAACATATGTCACCTTATTGCTGTGTAGACTCCCGAAGACAGCT	300
Db		252	CATAAGCTCTTAGSCAACATATGTCACCTTATTGCTGTGTAGACTCTCGAAGACCTGCA	311
OY		301	AACGATGTGAAAACTATTGGAACACTAATCTTCTTAAGAAAGTTAATACTACTAAATTT	360
Db		312	AACGATGTCAAACACTATTGGAACAACCACCTTCGAAAGAAGTTAANTGCTCCTCAGAT	371
OY		361	GTTCTTCGCCAAAAAGATTAAACAATAGTGTGAGAAATTGTAATAAGATTGAATTAATA	420
Db		372	CAGAAACAAGAGCAAGAGACTAGCAAAAGCCATGAAA--ATTACCGAACAACATATA	429
OY		421	AAACCTCAAGAGAGCAAGTATTTCTCAAGACAATGAGAATGTTTT-----	465
Db		430	AAACCTGCTGCTCGGACCTTCTCAAGGCCGCGCAATGATCATGTTCTTGTGGACGCG	489
OY		466	---ACAAACATATATGTAATTTTGGACGAGAGAGAACTTGACAAGAAATAATAA----	517
Db		490	AAAAGTTGTATTAATAAACCTATAGACAAAGATGAAGGTACACAGAAATATAAAGTTT	549
OY		518	----GTGAGAAACAAACCTCCAGATGCATCGATGAGAACACGTAGATCCATGSGTATAAT	573
Db		550	ACTGATGAGAAAGAAAAACCGAAGAAATGATGATGATGATGATGATGATGATGATGATG	609
OY		574	TTACTGAAAAATTGCATGACGATATTGTAAGAAAGATGAC	613
Db		610	TTATTAGCCAACAACATGAGATTGAAGACTTAGCTTAATG	649
RESULT_5				
AF146703				
LOCUS				
DEFINITION			Petunia integrifolia An2 protein (an2) mRNA, an2-S9 allele,	
ACCESSION			AF146703	
VERSION			AF146703	
KEYWORDS			complete cds.	
SOURCE			AF146703.1 GI:7673085	
ORGANISM			Petunia integrifolia	
REFERENCE			Bakayore; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS			Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
TITLE			1 (bases 1 to 865)	
JOURNAL			Molecular analysis of the anthocyanin2 gene of petunia and its role	
MEDLINE			Plant Cell 11 (8), 1433-1444 (1999)	
PUBMED			99380006	
REFERENCE			2 (bases 1 to 865)	
AUTHORS			Quattrocchio,F., Wing,J., van der Woude,K., Souer,E., de Vetten,N.,	
TITLE			Mol.J. and Koes,R.	
JOURNAL			Direct Submission	
FEATURES			Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan	
SOURCE			1087, Amsterdam 1081HV, Netherlands	
gene			location/Qualifiers	
			1..865	
			/organism="Petunia integrifolia"	
			/mol_type="mRNA"	
			/cultiivar="S9"	
			/db_xref="taxon:4103"	
			/chromosome="6"	
			/tissue_type="petal limb"	
			1..865	
			/gene="an2"	





QY	301	AACGATGGAAGAACTATTTGGAAACCTATCTTCAAGAGTTAAATACACTAAATY	360
Db	312	AAAGATGCAAAAATATTTGGACAACGCACTTGCAAGAAAGTTATTTCTTCATAT	371
QY	361	GTTCCTCGAGAAAGATTACAATTAAGTGTGAGAAATTAAGTACTAATTAATTA	420
Db	372	CAGAAACAAGAGAGCAACAAGCAGT-----AAATTAACGAGACAACTAATA	424
QY	421	AAACCTCAAGACGCAAGTATTTCTCAAGCAATGAAGAAATGT-----	465
Db	425	AAACGTGTCCTCGGACCTTCTCAAGCGCGCAATGATCATGTCTTGTGGAACGCG	484
QY	466	---ACAAACATTAATGTATTTTGGACGAGAGGAACCTTGCAGAAATTAATA----	517
Db	485	AAAAGTTGTAATAAAACACTATATGACACAGATGAAGGTGACACAGAAATATAAATTT	544
QY	518	----GTGAGAAACAACCTCCAGATGCAATGATGGAACAACGTAAAGATCCATGGTGATTAAT	573
Db	545	ACTGATGGAAGCAAAAACCGAACAATCGATAGATGATGGACTTCAATGGTGCGCCAAAT	604
QY	574	TTACTGAAAAATTCATGACGATTTTGAAGAGATGAG	613
Db	605	TTATTTAGCCAAACAACATTGAGATTGAGGCTTAAGCTAATG	644

LOCUS	AF146709	865 bp	mRNA	linear	PLN 01-MAY-2000
DEFINITION	Pectunia axillaris An2 truncated protein (an2) mRNA, an2-S7 allele,				
ACCESSION	AF146709				
VERSION	AF146709.1				
KEYWORDS	GI:7673097				
SOURCE	Pectunia axillaris				
ORGANISM	Pectunia axillaris				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamids; Solanales; Solanaceae; Pectunia.				
AUTHORS	1 (bases 1 to 865) Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N., Mol, J. and Koes, R.				
TITLE	Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color				
JOURNAL	Plant Cell 11 (8), 1433-1444 (1999)				
MEDLINE	99380006				
PUBMED	10449578				
REFERENCE	2 (bases 1 to 865) Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N., Mol, J. and Koes, R.				
AUTHORS	Direct Submission				
TITLE	Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan 1087, Amsterdam 1081HV, Netherlands				
JOURNAL	Location/Qualifiers				
FEATURES	1..865				
Source	/organism="Pectunia axillaris"				
	/mol_type="mRNA"				
	/cultivar="S7"				
	/db_xref="taxon:33119"				
	/chromosome="6"				
	/tissue_type="petal limb"				
gene	1..865				
	/gene="an2"				
	/allele="an2-S7"				
	12..374				
	/gene="an2"				
	/note="anthocyanin regulator; myb-related transcription factor"				
	/codon_start=1				
	/product="An2 truncated protein"				
	/protein_id="AAF6734.1"				
	/db_xref="GI:7673098"				
	/translation="MSTISNASTGSGYRKAMTTEEDDLIRETEIKYSGSKHLYVVRGALNRCSRLRLRLRYLRPHIKRGDFSLDEVDLILRLHLGLNRWSLLAGRLPGSTAND"				
CDS					

variation	VKQYMTDLRKKLIAPHD	372	/gene="a2"	/note="nonsense mutation"	
BASE COUNT	305 a	136 c	195 g	229 t	
ORIGIN					

Query Match	30.0%;	Score 247.8;	DB 8;	Length 865;
Best Local Similarity	62.1%;	Pred. No. 7.9e-39;		
Matches 448;	Conservative 0;	Mismatches 257;	Indels 16;	Gaps 3;

QY	1	ATGAAACAGTACATCTATGCTCTTCACTGGAGTGGAGAAAGGTTTCATGACCTGATGAA	60
Db	12	ATGAGTACTCTTAATGATCAACATCAGAGTAAGAAAGGTGATGACCGAGGAAGA	71
QY	61	GAATTTCTTCTAAGAAAATGATTGATAGTAGTGTAAGGAAAATGCAATCTTGTCCC	120
Db	72	GATCTTTTATGAGAGATGATCAATGAGAAAGTATGAGAAAGGAGTGGCAATGATGTTCCG	131
QY	121	ATTAAGACTGCTCTGAATTAATGATGCGAAAGTGTGATGTAAGTGGCTGAATTAATCTA	180
Db	132	GTTAGACTGCTCTGAATTAATGATGCGAAAGTGTGATGTAAGTGGCTGAATTAATCTA	191
QY	181	AGGCCACATATCAAGAGAGTGAATTTGAAACAAGATAGATGATCTCATTTGAGGCTT	240
Db	192	AGGCCACATATTAAGAGAGGACTTCTCTTTGATAGATGATCTCATTTTGGAGCTT	251
QY	241	CATAGCTCTTAGGCAACAGATGCTCATTAATGCTGGTAGACTTCCCGAAGACACT	300
Db	252	CATAGCTCTTAGGCAACAGATGCTCATTAATGCTGGTAGACTTCCCGAAGAACTGCA	311
QY	301	AACGATGTGAAAATCTTTGGAAACATTAATCTTTAAGAAAGTTAAATTAATCTAAAT	360
Db	312	AACGATGTCAAAAATCTTTGGAAACATTAATCTTTAAGAAAGTTAAATTTGCTCTCA	367
QY	361	GTTCTCGCGAAAAGATTAACTAATAGTGTGAGAAATTAATGATTAATGATTAATATA	420
Db	368	--TGATTAGAAACAAGAGACGAAAGCAAGCCCTGAAAATTAACGGAACACATATA	425
QY	421	AAACCTCAACGAGCGCAAGTATTTCTCAAGCAATAGAGATTTACAAACAATATGTA	480
Db	426	AAACCTGCTCTCGGACCTTCTCAAGCGGCAATGATGATGTTCTTGTGGAAAGGC	485
QY	481	ATTTTGAACGAGAGGAACATTTGCAGAGAAATTAATAGTGAAGAACTCCAGATGCA	540
Db	486	AAAAGTTGTAATAAAAAACATATAGCAAAATGAAAGTACACAGAAATTAATTAAGTTT	545
QY	541	TGCGATGACAAACGTAGATCC-----ATGCGATTAATTTCTGAAAATTCGATGA	593
Db	546	ACTGATAGAAAGCAAAACCGAGAAAGTCAATGATATGGAATCTCAATGCTGAGCMA	605
QY	594	CGATATTGAAGAGATGAAGGTTGTGTAATTAATATGAAAAAACAATAACAGTTGTT	653
Db	606	TTTATTAGCCAAACAATGATGATTAAGACTTAAGCTTAATGCTAATTCACCAATTTGTT	665
QY	654	ACATGAAGAAATATCACCACTTAATTAATTTGTAAGTGAAGTAACTCCATGCAACAGACA	713
Db	666	GCAATGAAGAAATATCACCACTTGTAAATAT---TGAGAGCAACTCATGCGAAGAGAGA	722
QY	714	A 714	
Db	723	A 723	

RESULT 8

AF146708

LOCUS

DEFINITION

Accession

VERSION

KEYWORDS

858 bp

mrna

linear

PLN 01-MAY-2000

Penutia axillaris An2 truncated protein (an2) mRNA, an2-S1 allele,

complete cds.

AF146708

AF146708.1 GI:7673095

SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCES  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

Petunia axillaris  
Petunia axillaris  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; lamiales; Solanales; Solanaceae; Petunia.  
1 (bases 1 to 858)  
Quattrocchio, F., Wang, J., van der Woude, K., Souer, E., de Vetten, N.,  
Mol, J., and Koes, R.  
Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color  
Plant Cell 11 (8), 1433-1444 (1999)  
99380006  
1049578  
2 (bases 1 to 858)  
Quattrocchio, F., Wang, J., van der Woude, K., Souer, E., de Vetten, N.,  
Mol, J., and Koes, R.  
Direct Submission  
Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan  
1087, Amsterdam 1081HV, Netherlands  
Location/Qualifiers  
1. 858  
/organism="Petunia axillaris"  
/mol\_type="mRNA"  
/cultivar="S1"  
/db\_xref="taxon:33119"  
/chromosome="6"  
/tissue\_type="petal limb"  
1. 858  
/gene="an2"  
/allele="an2-S1"  
12. 401  
/gene="an2"  
/note="anthocyanin regulator; myb-related transcription  
factor"  
/codon\_start=1  
/product="An2 truncated protein"  
/protein\_id="AA066733.1"  
/db\_xref="GI:7673096"  
/translation="MSTNSASTSGVRKAWTEBEDLLRPECIKYGGKHLVPAVG  
LNRCKRSCLRWLNYRPHIKRQDFSLDEVELLRHLKLGKMSLILAGRLPRTANN  
VKNYKNTLRKKLIVPHDQKESKTKP"  
391  
/gene="an2"  
/note="causes frameshift leading to truncated protein"  
/replace="n"  
BASE COUNT 304 a 134 c 194 g 226 t

variation

Query Match 29.6%; Score 244.4; DB 8; Length 858;  
Best Local Similarity 65.6%; Pred. No. 3,7e-38;  
Matches 420; Conservative 0; Mismatches 186; Indels 34; Gaps 3;

QY 1 ATGAACGATCATCTATGCTCTTCAATGGAGTGAAGAAAGTTCATGACTGATGAAGAA 60  
DB 12 ATAGATCACTCTATGATCATCAATCAGAGTAAGAAAGTGCATGACCGAGAGAA 71  
QY 61 GATTTCTTGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
DB 72 GATCTTTTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 131  
QY 121 ATAGAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
DB 132 GTTAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 191  
QY 181 AGGCCAATATCAAGAGAGTGAAGTTCGAAAGTTCGATGATGATGATGATGATGATGAT 240  
DB 192 AGGCCAATATCAAGAGAGTTCGATGATGATGATGATGATGATGATGATGATGATGAT 251  
QY 241 CATTAAGCTTGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 252 CATTAAGCTTGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 311  
QY 301 AACGATGAAAAAATAATGGAACATAATCTTCTTAAGAAAGTTAAATACCTAAATAT 360

DB 312 AACAAATGCAAAAATCAATTGGAAACAGCACTTCGAAAGAGTTAAATGTTCTCATGAT 371  
QY 361 GTTCTCGGAAAGATTAACAATAGTGGAAATAGTCTAGATGAATGAATATA 420  
DB 372 CAGAAACAGAGGCAAGCAAAAGCCATG-----AAATTAACGAAACAAATATA 424  
QY 421 AAACCTCAAGCAGCAAGTATTTCTCAAGCAGCAAGTAAAGATTT-----465  
DB 425 AAACCTGCTCTCGAACCTTCTCAAGCGCGCATATATCTTTCTTTGGAAACGC 484  
QY 466 ---ACAAACATTAATGTAATTTTGAAGAGAGAAACATTCGAGAAATATA-----517  
DB 485 AAAGTTGTAATTAATAACACTATAGAAGAAATGAAGTACACAGAAATATAAGTTT 544  
QY 518 ---GTGAGAAACAACTCCAGTGCATGATGATGATGATGATGATGATGATGATGATGAT 573  
DB 545 ACTGATGAGAGAGAAACCGGAAAGTTGATGATGATGATGATGATGATGATGATGATGAT 604  
QY 574 TTAATGCAAAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 613  
DB 605 TTATTTAGCCAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 644

RESULT 9  
AB073010 920 bp mRNA linear PLN 21-NOV-2002  
LOCUS  
DEFINITION  
Vitis labrusca x Vitis vinifera Vmlybdl-1 gene for myb-related  
transcription factor VIMYBdl-1, complete cds.  
AB073010  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Vitis labrusca x Vitis vinifera  
Vitis labrusca x Vitis vinifera  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; Vitaceae; Vitis.  
1  
Kobayashi, S., Ishimaru, M., Hiraoka, K. and Honda, C.  
Myb-related genes of the Kyoto grape (Vitis labruscana) regulate  
anthocyanin biosynthesis  
Planta 215 (6), 924-933 (2002)  
22242703  
12355152  
2 (bases 1 to 920)  
Kobayashi, S.  
Direct Submission  
Submitted (15-OCT-2001) Shozo Kobayashi, National Institute of  
Fruit Tree Science, Department of Grape and Persimmon Research,  
301-2, Akiitsu, Toyotsu-gun, Hiroshima 729-2494, Japan  
(E-mail:skobayashi@affrc.go.jp, Tel:81-846-45-4719,  
Fax:81-846-45-5370)  
29681  
Location/Qualifiers  
1. 920  
/organism="Vitis labrusca x Vitis vinifera"  
/mol\_type="mRNA"  
/cultivar="Kyoho"  
/db\_xref="taxon:105599"  
/tissue\_type="berry"  
/dev\_stage="mature"  
/cissue\_lib="Uni-ZAP Berry cDNA"  
1. 920  
/gene="Vimlybdl-1"  
15. 767  
/gene="Vimlybdl-1"  
/function="activate the Kyoho UFGT gene"  
/note="anthocyanin regulator"  
/codon\_start=1  
/product="myb-related transcription factor VIMYBdl-1"  
/protein\_id="BA07537.1"  
/db\_xref="GI:22266661"  
/translation="MESTLGVKGAAMTIOEDVLRKCIETKYGKHLVPLRAGLNRCS  
KSCRRLWNLKPDIKRGFALDEVLDLMIRLNLHNLGNRMSLILAGRLPRTADVKNYV





JOURNAL phenylpropanoid biosynthesis  
MEDLINE Plant Cell 12 (12), 2383-2394 (2000)  
PUBMED 21065949  
11148285  
REFERENCE 2 (bases 1 to 992)  
AUTHORS Borevitz, J.O., Xia, Y., Blount, J., Dixon, R.A. and Lamb, C.  
TITLE Direct Submission  
JOURNAL Submitted (29-NOV-2000) Plant Biology, Salk Institute, 10010 N  
Torrey Pines Rd, La Jolla, CA 92037, USA  
FEATURES  
source  
1. .992  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/cultivar="Columbia-0"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/map="near m412"  
/clone="BAC T27F4"  
1. .992  
/gene="PAP2"  
/note="synonyms: T27F4.14, MYB90"  
96. .845  
/gene="PAP2"  
/note="MYB transcription factor"  
/codon\_start=1  
/product="production of anthocyanin pigment 2 protein"  
/protein\_id="AA042002.1"  
/db\_xref="GI:11935173"  
/translation="MEGSKGLKRGAWTAEDSLRLCIDYGEKMHQVPLRAGINR  
CRKSCRLRWLNYLPSIKRGLSNDEVDLRLHLKLGKNSLJAGLPERTDVYN  
YMWLHLSKHESSCKSKMKKKNIISPTTPVKIGVFKRPSFSVNNKSHLNGLP  
EVDLPSCLGKKNVCENSITCNKDEKDFVNNLMGNNMLLEMLGENQADALIV  
PEATLHEHGKTLAPDVEQLMSLPDEIVEILD"

BASE COUNT 308 a 153 c 221 g 310 t  
ORIGIN

Query Match 22.5%; Score 186; DB 8; Length 992;  
Best Local Similarity 74.5%; Pred. No. 1.1e-26;  
Matches 234; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 19 TCTTCATTGGGAGTGAAGAAAGTTTCATGACGTGATGAAGAAGTTTCTTCTAAGAAA 78  
DB 105 TCGTCCAAAGGGTTGAGAAAGGTGATGACGTGATGAAGAAGTACTCTTGAGAGCTA 164  
QY 79 TGTATTGATTAAGTATGATGTAAGAAATGCGATTTTCCCAATGAAGCTGGCTGAAT 138  
DB 165 TGTATTGATTAAGTATGTAAGAAAGGCAATGCAATCAAGTTCTTTGAGAGCTGGCTAAT 224  
QY 139 AGATGTCGAAAGATTGATGAGTGGCTGAATATCTAAGGCCACATATCAAGAGA 198  
DB 225 CGATGCAAGAAAGTTTGAACCTAAGTGTGTAACCTATTGAAGCCAAAGTATCAAGAGA 284  
QY 199 GGTGACTTTGAACAAGATGATGATCTCATTTTGAAGCTTCATAAGCTTTAGGCAAC 258  
DB 285 GGAAGACTTGAACATGATGATGATCTCTTCTTGCCCTTCATTAAGCTTTAGGAAT 344  
QY 259 AGATGTCATCTTATGCTGTGAAGCTCCCGGAAGGACACTAAGCTGGAAGAAATAT 318  
DB 345 AGGTGGTCTTGAATGCTGTGATGCTCGTGGTGGACCGCTAATGATGTCAAAAATTAC 404  
QY 319 TGAACACTAATCT 332  
DB 405 TGAACACCATCT 418

RESULT 14  
AX358283 1033 bp DNA linear PAT 13-FEB-2002  
LOCUS  
DEFINITION Sequence 3 from Patent WO0200902.  
ACCESSION AX358283  
VERSION AX358283.1 GI:18674938  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1  
AUTHORS Borevitz, J., Xia, Y., Lamb, C.J. and Dixon, R.A.  
TITLE Regulation of anthocyanin pigment production  
JOURNAL Patent: WO 0200902-A 3 03-JAN-2002;  
THE SALK INSTITUTE FOR BIOLOGICAL STUDIES (US) ; THE SAMUEL ROBERTS  
NOBLE FOUNDATION, INC. (US)  
FEATURES  
source  
1. .1033  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"

BASE COUNT 334 a 161 c 226 g 312 t  
ORIGIN

Query Match 22.5%; Score 186; DB 6; Length 1033;  
Best Local Similarity 74.5%; Pred. No. 1.1e-26;  
Matches 234; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 19 TCTTCATTGGGAGTGAAGAAAGTTTCATGACGTGATGAAGAAGTTTCTTCTAAGAAA 78  
DB 122 TCGTCCAAAGGGTTGAGAAAGGTGATGACGTGATGAAGAAGTACTCTTGAGAGCTA 181  
QY 79 TGTATTGATTAAGTATGATGTAAGAAATGCGATTTTCCCAATGAAGCTGGCTGAAT 138  
DB 182 TGTATTGATTAAGTATGTAAGAAAGGCAATGCAATGTTCTTTGAGAGCTGGCTAAT 241  
QY 139 AGATGTCGAAAGATTGATGATGAGTGGCTGAATATCTAAGGCCACATATCAAGAGA 198  
DB 242 CGATGCAAGAAAGTTTGAACCTAAGTGTGTAACCTATTGAAGCCAAAGTATCAAGAGA 301  
QY 199 GGTGACTTTGAACAAGATGATGATCTCATTTTGAAGCTTCATAAGCTTTAGGCAAC 258  
DB 302 GGAAGACTTGAACATGATGATGATCTCTTCTTGCCCTTCATTAAGCTTTAGGAAT 361  
QY 259 AGATGTCATCTTATGCTGTGAAGCTCCCGGAAGGACACTAAGCTGGAAGAAATAT 318  
DB 362 AGGTGGTCTTGAATGCTGTGATGCTCGTGGTGGACCGCTAATGATGTCAAAAATTAC 421  
QY 319 TGAACACTAATCT 332  
DB 422 TGAACACCATCT 435

RESULT 15  
AF062915 1033 bp mRNA linear PLN 09-DEC-1998  
LOCUS  
DEFINITION Arabidopsis thaliana putative transcription factor (MYB90) mRNA,  
complete cds.  
ACCESSION AF062915  
VERSION AF062915.1 GI:3941521  
KEYWORDS  
SOURCE  
ORGANISM Arabidopsis thaliana (thale cress)  
Arabidopsids  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 1033)  
AUTHORS Kranz, H.D., Denekamp, M., Greco, R., Jin, H., Leyva, A., Meisener, R.C.,  
Petrioni, K., Urazzinski, A., Bevan, M., Martin, C., Smeekens, S.,  
Tonelli, C., Par-Arce, J. and Weisshaar, B.  
TITLE Towards functional characterisation of the members of the R2R3-MYB  
gene family from Arabidopsis thaliana  
JOURNAL Plant J. 16 (2), 263-276 (1998)  
MEDLINE 99056848  
PUBMED 9839469  
REFERENCE 2 (bases 1 to 1033)  
AUTHORS Greco, R., Petrioni, K. and Tonelli, C.  
TITLE Direct Submission  
JOURNAL Submitted (04-MAY-1998) Abt. Biochemie, MPI fur Zuchtungsforchung,







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 05:44:47 ; Search time 237 Seconds  
(without alignments)  
9396.778 Million cell updates/sec

Title: US-10-033-190-1

Sequence: 1 atgaacagatcattatgc.....tatgaatctactgtatcaa 825

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: N\_Geneseq\_19Jun03:\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	825	100.0	825	24	ABQ73046 Tomato anthocyanin
2	618	74.9	1012	24	ABQ73048 Tomato anthocyanin
3	189.6	23.0	762	22	ABQ5768 Arabidopsis thalia
4	189.6	23.0	762	24	ABK5362 Arabidopsis thalia
5	186	22.5	982	24	ABN98500 Arabidopsis thalia
6	186	22.5	1033	22	AAD5751 Arabidopsis thalia
7	186	22.5	1033	24	ABK5155 Arabidopsis thalia
8	186	22.5	1033	24	ABK14222 A. thaliana Produc

9	186	22.5	1033	25	AAD47497 Arabidopsis thalia
10	186	22.5	1043	22	AAF85191 Nucleotide sequenc
11	185.4	22.5	741	22	AAD05767 Arabidopsis thalia
12	185.4	22.5	741	24	ABK65363 Arabidopsis thalia
13	180.2	21.8	956	24	ABK14221 Arabidopsis thalia
14	179	21.7	747	24	ABK14829 Arabidopsis thalia
15	179	21.7	925	24	ABK65341 Arabidopsis thalia
16	170.6	20.7	390	25	ABX20395 Arabidopsis thalia
17	157.4	19.1	514	24	ABT08113 Arabidopsis thalia
18	154.8	18.8	774	21	AA42562 Arabidopsis thalia
19	154.8	18.8	774	24	ABK14831 Arabidopsis thalia
20	153.2	18.6	504	21	AAK56371 Arabidopsis thalia
21	153	18.5	933	22	AAK85190 Arabidopsis thalia
22	147	17.8	764	22	AAK80412 Arabidopsis thalia
23	147	17.8	791	22	AAK99077 Arabidopsis thalia
24	147	17.8	868	21	AAK49989 Arabidopsis thalia
25	147	17.8	871	21	AAK34284 Arabidopsis thalia
26	146.8	17.8	839	21	AAK51202 Arabidopsis thalia
27	144.4	17.5	1151	22	AAK90595 Arabidopsis thalia
28	143.8	17.4	1006	22	AAK90594 Arabidopsis thalia
29	143.8	17.4	1081	22	AAK90596 Arabidopsis thalia
30	143.4	17.4	612	24	ABK65317 Arabidopsis thalia
31	143	17.3	810	24	ABK14541 Arabidopsis thalia
32	143	17.3	839	21	AAK42479 Arabidopsis thalia
33	143	17.3	918	22	AAK06466 Arabidopsis thalia
34	139.6	16.9	1332	21	AAK40461 Arabidopsis thalia
35	138.6	16.8	660	24	ABK65183 Arabidopsis thalia
36	138	16.7	373	21	AAK57227 Arabidopsis thalia
37	138	16.7	960	24	ABK14533 Arabidopsis thalia
38	138	16.7	1280	21	AAK50981 Arabidopsis thalia
39	138	16.7	1310	22	AAK84746 Arabidopsis thalia
40	138	16.7	1329	21	AAK49321 Arabidopsis thalia
41	138	16.7	1392	21	AAK47037 Arabidopsis thalia
42	136.2	16.5	1082	22	AAK06651 Arabidopsis thalia
43	135.8	16.5	1236	24	ABT08114 Arabidopsis thalia
44	135.4	16.4	452	24	ABK93682 Arabidopsis thalia
45	134.6	16.3	1203	21	AAK33633 Arabidopsis thalia

## ALIGNMENTS

RESULT 1	ABQ73046 standard; cDNA; 825 BP.
ID	ABQ73046
AC	ABQ73046;
XX	24-SEP-2002 (first entry)
DT	Tomato anthocyanin 1 (ANT1) encoding cDNA SEQ ID NO.1.
DE	Tomato; ANT1; anthocyanin 1; Lycopersicon esculentum; plant; colour;
XX	gene; ss.
KW	Lycopersicon esculentum.
XX	OS
XX	Key
XX	Location/Qualifiers
XX	CDS
XX	1..825
XX	/*tag= a
XX	/product= "anthocyanin mutant (ANT1)"
XX	MO200255658-A2.
XX	18-JUL-2002.
XX	29-OCT-2001; 2001MO-US50638.
XX	30-OCT-2000; 2000US-244685P.
XX	(EXEL-) EXELIXIS PLANT SCI INC.
XX	Conners K, Mathews HV, Liu A;
XX	PI

```

XX MPI: 2002-557819/59.
DR P-PSDB; ABB81626.
XX
PT New isolated polynucleotide derived from tomato, useful for producing
PT an Anthocyanin 1 phenotype in plants, particularly for modifying e.g.
PT leaf color, flower color or fruit color in plants
XX
PS Claim 2; Page 39; 53bp; English.
XX
CC The present sequence encodes tomato anthocyanin 1 (ANT1). The ANT1
CC polynucleotide can be used for modifying the expression of a native
CC plant gene, particularly for producing an anthocyanin 1 phenotype in
CC plants, which is responsible for many red and blue colours in plants.
CC The polynucleotide is useful for modifying e.g. leaf colour, flower
CC colour or fruit colour in plants.
XX
SQ Sequence 825 BP; 309 A; 115 C; 175 G; 226 T; 0 other;

Query Match      100.0%; Score 825; DB 24; Length 825;
Best Local Similarity 100.0%; Pred. No. 5,1e-193;
Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACAGTACATCTATGCTTCTCATTTGGAGTGAAGAAAGTTTCATGAGTATGAAGAA 60
DB 1 ATGAACAGTACATCTATGCTTCTCATTTGGAGTGAAGAAAGTTTCATGAGTATGAAGAA 60
QY 61 GATTTCTTCTTAAGAAATGATATGATAGTGAAGAAATGGCACTTGTCC 120
DB 61 GATTTCTTCTTAAGAAATGATATGATAGTGAAGAAATGGCACTTGTCC 120
QY 61 GATTTCTTCTTAAGAAATGATATGATAGTGAAGAAATGGCACTTGTCC 120
DB 61 GATTTCTTCTTAAGAAATGATATGATAGTGAAGAAATGGCACTTGTCC 120
QY 121 ATAGAGCTGGTCTGAATAGATGTCGAAAGTTGATGAGTGGCTGATATCTA 180
DB 121 ATAGAGCTGGTCTGAATAGATGTCGAAAGTTGATGAGTGGCTGATATCTA 180
QY 121 ATAGAGCTGGTCTGAATAGATGTCGAAAGTTGATGAGTGGCTGATATCTA 180
DB 121 ATAGAGCTGGTCTGAATAGATGTCGAAAGTTGATGAGTGGCTGATATCTA 180
QY 181 AGGCCACATATCAAGAGAGTGTGACTTTGAACAGATGATGATCTCATTTTGAAGCTT 240
DB 181 AGGCCACATATCAAGAGAGTGTGACTTTGAACAGATGATGATCTCATTTTGAAGCTT 240
QY 181 AGGCCACATATCAAGAGAGTGTGACTTTGAACAGATGATGATCTCATTTTGAAGCTT 240
DB 181 AGGCCACATATCAAGAGAGTGTGACTTTGAACAGATGATGATCTCATTTTGAAGCTT 240
QY 241 CATAGCTCTTGAAGCAAGATGTCCTTATGCTGTGATGACTTCCCGAAGGACAGCT 300
DB 241 CATAGCTCTTGAAGCAAGATGTCCTTATGCTGTGATGACTTCCCGAAGGACAGCT 300
QY 241 CATAGCTCTTGAAGCAAGATGTCCTTATGCTGTGATGACTTCCCGAAGGACAGCT 300
DB 241 CATAGCTCTTGAAGCAAGATGTCCTTATGCTGTGATGACTTCCCGAAGGACAGCT 300
QY 301 AACGATGGAAGAACTATTTGGAACACTATCTTCTAAGAAAGTTAAATCTACTAAAT 360
DB 301 AACGATGGAAGAACTATTTGGAACACTATCTTCTAAGAAAGTTAAATCTACTAAAT 360
QY 301 AACGATGGAAGAACTATTTGGAACACTATCTTCTAAGAAAGTTAAATCTACTAAAT 360
DB 301 AACGATGGAAGAACTATTTGGAACACTATCTTCTAAGAAAGTTAAATCTACTAAAT 360
QY 361 GTTCTCTGCGAAAGATTAAACATTAAGTGTGAGAAATTAGTACTTAAGATTGAATTTA 420
DB 361 GTTCTCTGCGAAAGATTAAACATTAAGTGTGAGAAATTAGTACTTAAGATTGAATTTA 420
QY 361 GTTCTCTGCGAAAGATTAAACATTAAGTGTGAGAAATTAGTACTTAAGATTGAATTTA 420
DB 361 GTTCTCTGCGAAAGATTAAACATTAAGTGTGAGAAATTAGTACTTAAGATTGAATTTA 420
QY 421 AAACCTCAAGACCCAGTATTTCTCAAGCACAATGAAGATGTTACCAAAATTAATGTA 480
DB 421 AAACCTCAAGACCCAGTATTTCTCAAGCACAATGAAGATGTTACCAAAATTAATGTA 480
QY 421 AAACCTCAAGACCCAGTATTTCTCAAGCACAATGAAGATGTTACCAAAATTAATGTA 480
DB 421 AAACCTCAAGACCCAGTATTTCTCAAGCACAATGAAGATGTTACCAAAATTAATGTA 480
QY 481 ATTTTGAACGAGGAACATTTGCAAGAAATTAATTAAGTGAAGAAACCACTCCAGATGCA 540
DB 481 ATTTTGAACGAGGAACATTTGCAAGAAATTAATTAAGTGAAGAAACCACTCCAGATGCA 540
QY 481 ATTTTGAACGAGGAACATTTGCAAGAAATTAATTAAGTGAAGAAACCACTCCAGATGCA 540
DB 481 ATTTTGAACGAGGAACATTTGCAAGAAATTAATTAAGTGAAGAAACCACTCCAGATGCA 540
QY 541 TCGATGCAACGATAGATCCATGCTGATTAATTAATTTACTGAAAAATTGCAATGATATT 600
DB 541 TCGATGCAACGATAGATCCATGCTGATTAATTAATTTACTGAAAAATTGCAATGATATT 600
QY 541 TCGATGCAACGATAGATCCATGCTGATTAATTAATTTACTGAAAAATTGCAATGATATT 600
DB 541 TCGATGCAACGATAGATCCATGCTGATTAATTAATTTACTGAAAAATTGCAATGATATT 600
QY 601 GAAAGAAATGAAGAGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
DB 601 GAAAGAAATGAAGAGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
QY 601 GAAAGAAATGAAGAGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
DB 601 GAAAGAAATGAAGAGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
QY 661 GAAATATCACCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
DB 661 GAAATATCACCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
QY 661 GAAATATCACCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
DB 661 GAAATATCACCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
QY 721 CATGAAATTTGGGGTGAATTTTCTTAAATTACACCCATGCAACAGAGTAAAT 780
DB 721 CATGAAATTTGGGGTGAATTTTCTTAAATTACACCCATGCAACAGAGTAAAT 780

```

```

DB 721 CATGAAATTTGGGGTGAATTTTCTTAAATTACACCCATGCAACAGAGTAAAT 780
QY 781 GATGATTTTCTGCTGAATTAATGACTTATGGAATCTTATGATTA 825
DB 781 GATGATTTTCTGCTGAATTAATGACTTATGGAATCTTATGATTA 825

RESULT 2
ABQ73048
ID ABQ73048 standard; cDNA; 1012 BP.
XX
AC ABQ73048;
XX
DT 24-SEP-2002 (first entry)
XX
DE Tomato anthocyanin 1 (ANT1) genomic fragment SEQ ID NO:4.
XX
KM Tomato; ANT1; anthocyanin 1; Lycopersicon esculentum; plant; colour;
KW gene; de.
XX
OS Lycopersicon esculentum.
XX
PN NO200255658-A2.
XX
PD 18-JUL-2002.
XX
PF 29-OCT-2001; 2001MO-US50638.
XX
PR 30-OCT-2000; 2000US-244685P.
XX
PA (EXEL-) EXELIXIS PLANT SCI INC.
XX
PI Connors K, Mathews HV, Liu A;
XX
DR MPI: 2002-557819/59.
XX
PT New isolated polynucleotide derived from tomato, useful for producing
PT an Anthocyanin 1 phenotype in plants, particularly for modifying e.g.
PT leaf color, flower color or fruit color in plants
XX
PS Example 3; Page 46-47; 53bp; English.
XX
CC The present invention describes tomato anthocyanin 1 (ANT1). The ANT1
CC polynucleotide can be used for modifying the expression of a native
CC plant gene, particularly for producing an anthocyanin 1 phenotype in
CC plants, which is responsible for many red and blue colours in plants.
CC The polynucleotide is useful for modifying e.g. leaf colour, flower
CC colour or fruit colour in plants. The present sequence represents an
CC ANT1 genomic DNA fragment which is used in an example from the present
CC invention.
XX
SQ Sequence 1012 BP; 364 A; 138 C; 201 G; 309 T; 0 other;

Query Match      74.9%; Score 618; DB 24; Length 1012;
Best Local Similarity 81.5%; Pred. No. 3.9e-142;
Matches 825; Conservative 0; Mismatches 0; Indels 187; Gaps 2;

QY 1 ATGAACAGTACATCTATGCTTCTCATTTGGAGTGAAGAAAGTTTCATGAGTATGAAGAA 60
DB 1 ATGAACAGTACATCTATGCTTCTCATTTGGAGTGAAGAAAGTTTCATGAGTATGAAGAA 60
QY 61 GATTTCTTCTTAAGAAATGATATGATAGTGAAGAAATGGCACTTGTCC 120
DB 61 GATTTCTTCTTAAGAAATGATATGATAGTGAAGAAATGGCACTTGTCC 120
QY 61 GATTTCTTCTTAAGAAATGATATGATAGTGAAGAAATGGCACTTGTCC 120
DB 61 GATTTCTTCTTAAGAAATGATATGATAGTGAAGAAATGGCACTTGTCC 120
QY 121 ATAGAGCTGGTCTGAATAGATGTCGAAAGTTGATGAGTGGCTGATATCTA 180
DB 121 ATAGAGCTGGTCTGAATAGATGTCGAAAGTTGATGAGTGGCTGATATCTA 180
QY 121 ATAGAGCTGGTCTGAATAGATGTCGAAAGTTGATGAGTGGCTGATATCTA 180
DB 121 ATAGAGCTGGTCTGAATAGATGTCGAAAGTTGATGAGTGGCTGATATCTA 180
QY 130 -----GGTGAATGATGT 144
DB 130 -----GGTGAATGATGT 144
QY 181 TTTTATTGACGTTATTACGAATATCATGAAATATGATGTCAGAGTCTGAATAGATGT 240
DB 181 TTTTATTGACGTTATTACGAATATCATGAAATATGATGTCAGAGTCTGAATAGATGT 240

```

Oy		145	CGGAAAAGTTGATGATTGAGTGGCTGCAGTAATTACTAAGGCCACATATCAAGAAGGTGAC	204
Db		241	CGGAAAAGTTGATGATTGAGTGGCTGCAGTAATTACTAAGGCCACATATCAAGAAGGTGAC	300
Oy		205	TTTGAAACAAGATGAAGTGCATCTCATTTTGAAGGCTTCATAAGCTCTTAGGCAAC-----	258
Db		301	TTTGAACAAGATGAAGTGCATCTCATTTTGAAGGCTTCATAAGCTCTTAGGCAACAAGCAT	360
Oy		259	-----	258
Db		361	GCAAGTTATGTTTTGCACAAAAATTTGATTAGTATATATATATATACGTGTGACTATTTTC	420
Oy		259	-----AGATGGTACTTATTTGCTGGTGAAGACTCCCGGAAG	293
Db		421	ATCTTAAAGTTACGTTATTTTACGTGATGATGGTACATTATGCTGGTAGACTTCCCGGAAG	480
Oy		294	GACAGCTAACGATGTGTA AAAA ACTAATTGGAACA CTAATCTCTCTAAGGAAGTTAATACTAC	353
Db		481	GACAGCTAACGATGTGTA AAAA ACTAATTGGAACA CTAATCTCTCTAAGGAAGTTAATACTAC	540
Oy		354	TAAATTTGTTCTCCGCGAAAAGATTACATAAAGTGTGAGAAAATTAGTACTAAGATTGA	413
Db		541	TAAATTTGTTCTCCGCGAAAAGATTACATAAAGTGTGAGAAAATTAGTACTAAGATTGA	600
Oy		414	AATTTPAAAACCTCAACGACGCAAGTAATTTCTCAACACAAATGAAGTTTCAAACAA	473
Db		601	AATTTPAAAACCTCAACGACGCAAGTAATTTCTCAACACAAATGAAGTTTCAAACAA	660
Oy		474	TAATGTAAATTTTGGACAGSAGSAAACATTGGCAAGAAATTAATAAGTGAGAAA CMAACTCC	533
Db		661	TAATGTAAATTTTGGACAGSAGSAAACATTGGCAAGAAATTAATAAGTGAGAAA CMAACTCC	720
Oy		534	AGATSCATCATGATGACAAACGTAGATCCATGGTGGATPAAATTNACTGGA AAATTGCAATGA	593
Db		721	AGATSCATCATGATGACAAACGTAGATCCATGGTGGATPAAATTNACTGGA AAATTGCAATGA	780
Oy		594	CGATTTTGAAGAAGATGAAGAGTTGTATTTAATTATGAAAAAAACCTAACAGTTTGT	653
Db		781	CGATTTTGAAGAAGATGAAGAGTTGTATTTAATTATGAAAAAAACCTAACAGTTTGT	840
Oy		654	ACATGAAGA AATATATCACCACCACTTAAATATTTGGTGAAGGTAAC TCCATGCAACAAAGACA	713
Db		841	ACATGAAGA AATATATCACCACCACTTAAATATTTGGTGAAGGTAAC TCCATGCAACAAAGACA	900
Oy		714	AATTAAGTCATGAAAATTTGGGGTGAATTTTCTCTTAATTTAACACCCCATGCAACAAGAGT	773
Db		901	AATTAAGTCATGAAAATTTGGGGTGAATTTTCTCTTAATTTAACACCCCATGCAACAAGAGT	960
Oy		774	ACAAATATGATGATTTTCTGCTGAAATTAAGTACTTATGGAATCTACTGATTAA	825
Db		961	ACAAATATGATGATTTTCTGCTGAAATTAAGTACTTATGGAATCTACTGATTAA	1012

RESULT\_3  
AD05768 ID AD05768 standard; cDNA; 762 BP.  
XX AC AAD05768;  
XX DT 31-JUL-2001 (first entry)  
XX DE Arabidopsis thaliana transcription factor homologue G2421 cDNA.  
XX KW Transcription factor; trait modification; seed characteristic;  
XX KM structural characteristic; developmental characteristic; gene therapy;  
XX OS agricultural biotechnology; ss.  
XX Arachidopsis thaliana.  
XX FH Key Location/Qualifiers  
XX CDS 1..630  
XX FT /\*tag= a "Transcription factor homologue G2421"  
FT /product= "Transcription factor homologue G2421"

[illegible]



PA (RAME/) RAMEKA J G.  
PA (PAGE/) PAGE A.  
PA (MATH/) MATHW A V.  
PA (LEDF/) LEDFORD B L.  
PA (MOES/) MOESSNER J P.  
PA (HARS/) HARS W D.  
PA (GARC/) GARCIA C A.  
PA (KRICK/) KRICKER M.  
PA (SLAT/) SLATER T.  
PA (DAVI/) DAVIS K R.  
PA (ALLE/) ALLEN K.  
PA (HOFF/) HOFFMAN N.  
PA (HURB/) HURBAN P.  
PI Goriach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;  
PI Rameka JG, Page A, Mathew AV, Ledford BL, Moesner JP, Haas WD;  
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
PI Hurban P;  
DR WPI; 2002-400781/43.  
XX  
XX New Arabidopsis thaliana nucleic acid for identifying homologous genes,  
PT producing compositions that modulate the expression or function of its  
PT encoded protein, and mapping functional regions of protein -  
XX  
XX Claim 1; SEQ ID NO 268; 49pp + Sequence Listing; English.  
XX  
XX The invention relates to an Arabidopsis thaliana nucleic acid (I)  
XX comprising a sequence capable of hybridizing under stringent conditions  
XX to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),  
XX given in the specification or its fragment. A polypeptide (II) encoded by  
XX (I), a transgenic plant (III) comprising an exogenous nucleic acid or a  
XX genetically modified cell (IV) comprising an exogenous nucleic acid, is  
XX useful for screening a candidate agent for its biological effect. (I) is  
XX useful in identifying homologous or related genes, in producing  
XX compositions that modulate the expression or function of its encoded  
XX protein, mapping functional regions of the protein and in studying  
XX associated physiological pathways. (I) is also useful for the genetic  
XX manipulation of cells, particularly plant cells. (I) is also useful in  
XX screening assays of various plant strains to determine the strains that  
XX are best capable of withstanding a particular disease or environmental  
XX stress. (II) and (III) are useful for screening of biologically active  
XX agents, e.g. fungicides, insecticides, etc., for elucidating biochemical  
XX pathways. The screened agents are useful in improved methods of treating  
XX crops to prevent or treat disease. (II) are also useful in screening  
XX programs to identify agents that mimic or enhance the action of tolerance  
XX factors. Such agents are useful in improved methods of treating crops to  
XX enhance their tolerance to environmental stress. (I) is also useful  
XX for enhancing or inhibiting production of a biosynthetic product in a  
XX plant. (III) is useful for identifying other mediators that may induce  
XX expression of proteins of interest, for establishing the extent to which  
XX any specific insect and/or pathogen is responsible for damage to a  
XX particular plant, for identifying other mediators that enhance or induce  
XX tolerance to environmental stress, for identifying factors involved in  
XX biosynthetic pathways of nutritional, commercial, or medicinal value and  
XX for identifying productions of nutritional, commercial or medicinal  
XX value. (IV) is useful in the study of genetic function and regulation,  
XX for alteration of the cellular metabolism and for screening compounds  
XX that may affect the biological function of the gene or gene products.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from USPTO  
XX at seqdata.uspto.gov/sequence.html?docid=999990770445.  
XX  
SQ Sequence 982 BP; 303 A; 215 C; 140 G; 322 T; 2 other;

Query Match 22.5%; Score 186; DB 24; Length 982;  
Best Local Similarity 74.5%; Pred. No. 5.3e-36;  
Matches 233; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

OY 19 TCTTCAATGGAGTGAAGAAAGTTTCATGACATGATGAAGAAAGATTTCTTCTTAAGAAA 78  
DB 911 TCGTCAAAAGGTGATGAAGAAAGTGACATGACTGTGAAGAAAGATGCTCTTGAAGCTA 852

OY 79 TGTATTGATTAAGTANGTGAAGAAAATGCGATCTTGTTCCATTAAGAGCTGTCGAAT 138  
DB 851 TGTATTGATTAAGTANGTGAAGAAAATGCGATCAAGTTCTTTGAGAGCTGGCTAAAT 792  
\* OY 139 AGATGTCGAAAAGTTGTAGATTGAGTGGCTGAAATTAATCTAAGGCCAATTAAGAGA 198  
DB 791 CGATGCAAGAAAGGTGTGACTAAGATGATGTGAATTAATTTGAAGCAAGTATCAAGAGA 732  
OY 199 GGTGACTTGAACAAGATGAAGTGAATCTCATTTGAGGCTTCATTAAGCTTTAGGCAAC 258  
DB 731 GGAAGACTTAAGCATATATAGATTGATCTTCTTCCGCTTCATTAAGCTTCAGGAAT 672  
OY 259 AGATGTCACCTTAATTCCTGATGACTTCCCGAAGAGACATTAAGATGAAAACATAT 318  
DB 671 AGGTGATCCTTGATTTGCTGATGATTTGCTGATGCTGCGACCGCTAATGATGCAAAATTAAC 612  
OY 319 TGAACACTTAATCT 332  
DB 611 TGAACACCCATCT 598  
RESULT 6  
AAB05751  
ID AAB05751 standard; cDNA; 1033 BP.  
XX  
XX AAB05751;  
AC  
XX  
DT 31-JUL-2001 (first entry)  
XX  
XX Arabidopsis thaliana transcription factor G663 cDNA.  
DE  
XX Transcription factor; trait modification; seed characteristic;  
KW structural characteristic; developmental characteristic; gene therapy;  
KW agricultural biotechnology; ss.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 113..862  
FT /\*tag = a  
FT /product = "Transcription factor G663"  
XX  
XX MO200135727-A1.  
XX  
PD 25-MAY-2001.  
XX  
XX 14-NOV-2000; 2000MO-US31457.  
XX  
XX 17-NOV-1999; 99US-0166228.  
PR 17-APR-2000; 2000US-0197899.  
PR 22-AUG-2000; 2000US-0227439.  
XX  
XX (MENDEL) MENDEL BIOTECHNOLOGY INC.  
PA (REUB/) REUBER L.  
PA (CREE/) CREELMAN R.  
PA (PILG/) PILGRIM M.  
PA (RIEC/) RIECHMAN J L.  
PA (JIAN/) JIANG C.  
PA (YUGG/) YU G.  
PA (PINE/) PINEDA O.  
PA (HEAR/) HEARD J.  
XX  
PI Reuber L, Creelman R, Pilgrim M, Riechmann JL, Jiang C, Yu G;  
PI Pineda O, Heard J;  
XX  
XX WPI; 2001-335979/35.  
XX P-PSDB; AAE01916.  
XX  
PT Nucleic acids encoding plant transcription factor polypeptides, useful  
PT for altering the developmental and structural characteristics of  
PT plants, e.g. corn, potato and cotton plants -  
XX  
PS Claim 4; Page 69-70; 133pp; English.



Db	302	GGAAGACTTGAATGATGAAGTTGATCTTCTTCCTTCGCTTCATTAAGCTCTAGAAT	3
Qy	259	AGATGCTCACTTATTTGCTGTAGACTTCCCGAAGACAGCTAAGATGTAAAACTAT	311
Db	362	AGGTGGTCTTGATTTGCTGTGATTCGATTCCTGTGCGACCGCTAATGATGTCAAAATTAC	422
Qy	319	TGGAACTAATCT	332
Db	422	TGGAACTAATCT	435
RESULT 8			
ID	ABK14222	standard; cDNA; 1033 BP.	
AC	ABK14222;		
DT	08-MAY-2002	(first entry)	
DE	A. thaliana	Production of anthocyanin pigment 2 (PAP2) cDNA.	
KW	PAP2; production of anthocyanin pigment; MYB-like transcription factor;		
KW	transgenic; plant; phenylpropanoid; stress; light stress; water stress;		
KW	pH stress; temperature stress; heavy metal stress; pathogen attack;		
KW	infection; wounding; nutrient deficiency; herbivory; plant colouration;		
KW	stress sensitisation; gene; ss; chromosome 1.		
XX	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		
XX			
XX	Key	Location/Qualifiers	
FT	CDS	113..862	
FT		/*tag= a	
FT		/product= "PAP2 protein"	
FT		/transl_except= (pos:500..502, aa:Ser)	
FT		/transl_except= (pos:593..595, aa:Xaa)	
FT		/transl_except= (pos:680..682, aa:Xaa)	
FT		/transl_except= (pos:755..757, aa:Xaa)	
FT		/transl_except= (pos:842..844, aa:Lys)	
XX			
XX	MO200200902-A2.		
XX			
PD	03-JAN-2002.		
XX			
PF	21-JUN-2001; 2001WO-US19734.		
XX			
PR	23-JUN-2000; 2000US-0603244.		
PR	05-JUL-2000; 2000US-0610185.		
XX			
PA	(SALK ) SALK INST BIOLOGICAL STUDIES.		
PA	(ROBE-) ROBERTS NOBLE FOUND INC SAMMUEL.		
XX			
PI	Borevitz J, Xia Y, Lamb CJ, Dixon RA;		
XX			
DR	WPI; 2002-164443/21.		
DR	P-PSDB; AAU075735.		
XX			
PT	New isolated PAP1 or PAP2 gene, useful for increasing pigmentation in		
PT	plants, as reporter genes for analysing expression pattern of promoter		
PT	of interest, and to increase flux through phenylpropanoid pathway		
PS	Claim 5; Page 27; 29pp; English.		
XX			
XX	This invention relates to the nucleotide and protein sequences of novel		
CC	Arabidopsis thaliana production of anthocyanin pigment (PAP) genes PAP1		
CC	or PAP2. The genes for both PAP1 and PAP2 have been mapped to chromosome		
CC	1. PAP1 and PAP2 are MYB-like transcription factors that regulate the		
CC	production of anthocyanin pigment. The nucleotide sequence of the		
CC	invention is useful for screening for PAP gene expression, it is also		
CC	useful for detecting promoter activity within a plant cell. The		
CC	nucleotide sequence is also useful for screening for plant cell (e.g.,		
CC	A. thaliana) transformation. The transformed plant cells are then used		
CC	to produce transgenic plants with tissue-specific altered colour		
CC	expression. The nucleotide sequence is also useful for enhancing		

XX	CC	accumulation of phenylpropanoid product in a plant. A transgenic
CC	plant is useful for detecting a stress condition such as light	
CC	stress, water stress, pH stress, temperature stress, heavy metal	
CC	stress, pathogen attack or infection, wounding, nutrient deficiency,	
CC	herbivory, or abnormal hormone levels in a plant. PAP1 or PAP2	
CC	overexpression can also be useful in developing or enhancing desirable	
CC	colouration in ornamental plants for commercial sales, and for	
CC	sensitising plants to stress. Introduction of PAP1 or PAP2 sensitises	
CC	plants to respond more quickly to a stress condition in ways that are	
CC	physiologically significant, as well as in ways of ornamental	
CC	significance. The present sequence represents the A. thaliana PAP2 cDNA	
CC	sequence of the invention.	
XX		
SO	Sequence 1033 BP; 334 A; 161 C; 226 G; 312 T; 0 other;	
	Query Match 22.5%; Score 186; DB:24; Length 1033;	
	Best Local Similarity 74.5%; Pred. No. 5.4e-36;	
	Matches 234; Conservative 0; Mismatches 80; Indels 0; Gaps 0;	
QY	19 TCTTCATTGGGAGTGGAGAAAGTTTCATGCATGATGAGAAAGTTTCTTCTAAGAAA 78	
DB	122 TCGTCCAAAGGGTTGAGAAAGGTCATGAGCTGCGAAGAGATAGTCTTGGGCTA 181	
QY	79 TGTATTGATAGTATGAGGAAAGAAATGCGATCTTGTTCCCATAAAGCTGCTGAT 138	
DB	182 TGTATTGATAGTATGAGGAAAGGCAATGGCATCAAGTCTTTGAGAGCTGGGCTAAT 241	
QY	139 AGATGTGCGAAAAGTTGTAGATTGAGGTGCGCTGAATTATCTAAGGCCCATATCAAGAGA 198	
DB	242 CGATGCGAAAAGATTTGTAGCTTAAGATGGTTGAACATATTGAAAGCCAAAGTATCAAGAGA 301	
QY	199 GGTGACTTTGAAACAAGATGAAGTGACATCTATTTTGAGGCTTCATPAAGCTTTAGGCAC 258	
DB	302 GGAAGACTTTCAGATGATGAAGTTGATCTTCTCTTGGCCTTCATPAAGCTTTAGGAAAT 361	
QY	259 AGATGTGCTACTTATGCTGTGAGACTTCCCGGAGAGACAGTTACGATGTGAAAACTAT 318	
DB	362 AGGTGTGCTTGTGATGCTGTGCTGATGCTGCTGCGACCGCTAATGATGTCAAAAATTAC 421	
QY	319 TGGAACTACTAATCT 332	
DB	422 TGGAACTACTAATCT 435	
	RESULT 9	
	AAD47497	
ID	AAD47497 standard; DNA; 1033 BP.	
XX		
AC	AAD47497;	
XX		
DT	24-FEB-2003 (first entry)	
XX		
DE	Arabidopsis thaliana G663 transcription factor DNA.	
XX		
KW	Transcription factor; metabolite pathway; terpenoid; limonene synthase;	
KW	alcaloid pathway gene; taxadiene synthase; biological pathway; freezing;	
KW	abiotic stress; cold; drought; heat; nutrient deficiency; biotic stress;	
KW	infection; developmental pathway; flowering; root development; TDS; LS;	
KW	transgenic; transgenic plant; gene; ds.	
XX		
OS	Arabidopsis thaliana.	
XX		
FH	Key Location/Qualifiers	
FT	CDS 113..862	
FT	FT /tag=a	
FT	/product= "Arabidopsis thaliana G663	
XX	transcription factor protein"	
XX		
PN	W0200274917-A2.	
XX		
DD	26-SEP-2002.	
XX		
PF	15-MAR-2002; 2002WO-US07999.	



Query	Subject	Score	E-value	Length	Identical	Mismatches	Gaps
QY	19 TCTTCATTCGGAGTGAAGAAAGGTCATGACCTGATGAGAAAGATTTCTCTTAAGAAAA	78					
DB	122 TCGTCGCAAGAGGTTGAGAAAGGTGATGACGTGCTGAAGAAAGATGTCCTTTGAGGCTA	181					
QY	79 TGTATTGATTAAGTATGCTGTAAGAAATGGCATCTTGTTCCTAATAGCTGGTCTGAAT	138					
DB	182 TGTATTGATTAAGTATGAGAAAGGCAATATGCAATTCCTTTGAGACCTGGGCTTAAT	241					
QY	139 AGATGTGCGAAAGTTGTAGATTGAGGTGGCTGTAATTAATCTTAAGGACACATATCAAGAGA	198					
DB	242 CGATGACGAAAGAGTTGTAACTTAAGTGGTTAACTAATTTGAAGCAAGATCAAGAGA	301					
QY	199 GGTGACCTTTGAACAGATGAAGTGAATCTCATTTTGAAGGCTTCATTAAGCTCTTAGCAAC	258					
DB	302 GGAAGACTTGAACAATGATGAAGTTGATCTTCTTCCTGCTTATTAAGCTCTTAGGAAT	361					
QY	259 AGATGTGACTTATTTGCTGTGAACCTTCCGGAAGACAGCTTAACAGATGTGAAGAACTAT	318					
DB	362 AGGATGCTTCATGATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	421					
QY	319 TGGAACTATATCT 332						
DB	422 TGGAACTATATCT 435						

XX	AA85191;
AC	
XX	
DT	09-JUL-2001 (first entry)
XX	
DE	Nucleotide sequence of MYB transcription factor AtMYB90.
XX	
KW	Stress tolerance; mylolaostosis transcription factor; heat tolerance;
KM	MYB transcripition factor; AtMYB60; AtMYB764; AtMYB75; transgenic plant;
XX	phenylpropanoid biosynthesis; ss.
OS	
XX	Arabidopsis sp.
FH	
FT	Key Location/Qualifiers
FT	CDS 113..862
FT	/tag= a
FT	/product= "MYB transcription factor"
PN	
XX	MO200132002-A1.
PD	
XX	10-MAY-2001.
PF	
XX	06-NOV-2000; 2000WO-US30503.
PR	
XX	05-NOV-1999; 99US-0163579.
PR	23-OCT-2000; 2000US-0693855.
XX	
PA	(BADT ) BASF CORP.
PI	
XX	Tonell1 C;
DR	
XX	WP1; 2001-316365/33.
DR	P-PSDB; AAB68356.
PT	
XX	Nucleic acid molecules encoding plant stress tolerance-related
PT	mylolaostosis transcription factors for increasing stress resistance of
PT	a crop in a field -
XX	
PS	Claim 7; Fig 4; 92pp; English.
CC	
CC	The present sequence encodes a plant stress tolerance-related
CC	AtMYBlaostosis (MYB) transcription factor. The specification describes
CC	AtMYB60, AtMYB764, AtMYB75 and AtMYB90. MYB polypeptides are useful for
CC	enhancing a plant's tolerance or sensitivity to stress e.g. salt,
CC	drought, cold and heat tolerance. It is also useful for producing a
CC	transgenic plant with enhanced stress tolerance or stress sensitivity.
CC	MYB polypeptides are also useful for increasing the production of
CC	products of the phenylpropanoid biosynthesis pathway, e.g. stilbenes,
CC	flavonoids, lignins, salicylic acid, anthocyanins, and phenolic
CC	derivatives.
XX	
SQ	Sequence 1043 BP; 334 A; 166 C; 231 G; 312 T; 0 other;
BQ	
Match	22.5%; Score 186; DB 22; Length 1043;
Match	Local Similarity 74.5%; Pred. No. 5.4e-36;
Matches	234; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
OY	19 TCTTCATTGGAGTGAGAAAAGTTCAATGACTGATGAAGAAGATTTCTTCAAGAAAA 78
Db	122 TCGTCCAAGGTTTGAAGAAAGTGATGAGACTGCTGMAAAGATATCTCTTGAGGCTTA 181
OY	79 TGATTTGATTAAGTATGTTGGAAGAAATGCGATCTTGTGCCATTAAGAGCTGTGTAAT 138
Db	182 TGTATTTGATAAGTATGGAAGAAAGGCAATGGCATCAAGTTCCTTTGAGAGCTGGGCTTAAT 241
OY	139 AGATGTGGGAAGAGTTGTAGATGAGGTGGCTGAATTATCTTAAGGCCCATATCAAGAGA 198
Db	242 CGAATGCAGAAAGAGTTGATGACTAAGATGGTTGAATCTTAAGAACCAAGATCAAGAGA 301
OY	199 GGTAAGCTTTGAACAAGATGAAGTGCATCTCAATTTTGAAGCTTCAATTAAGCTTGAAGC 258
Db	302 GGAAGACTTACATGATGATGAAGTGTATCTTCTGCGCCTTCAATTAAGCTTGAAGAAAT 361

QY 259 AGATGCTACTTATGCTGTGAGACTTCCGGAAGACAGCTTACGATGTGAAAACTAT 318  
DB 362 AGGTGGCTCTTGAATGCTGTGCTGATGCTGTGCGACCGCTAATGATGTCAAAAATTAC 421  
QY 319 TGGAAACACTAATCT 332  
DB 422 TGGAAACACCATCT 435

RESULT 11  
AAD05767  
ID AAD05767 standard; cDNA; 741 BP.  
AC AAD05767;  
DT 31-JUL-2001 (first entry)  
XX Arabidopsis thaliana transcription factor homologue G2422 cDNA.  
XX  
XX Transcription factor; trait modification; seed characteristic;  
XX structural characteristic; developmental characteristic; gene therapy;  
XX agricultural biotechnology; ss.  
XX Arabidopsis thaliana.  
XX  
XX Key Location/Qualifiers  
FH CDS 1..741  
FT /\*tag= a  
FT /product= "Transcription factor homologue G2422"  
XX  
XX WO200135727-A1.  
XX  
XX 25-MAY-2001.  
XX  
XX 14-NOV-2000; 2000MO-US31457.  
XX  
XX 17-NOV-1999; 99US-0166228.  
XX 17-APR-2000; 2000US-0197899.  
XX 22-AUG-2000; 2000US-0227439.  
XX  
XX (MENDEL) MENDEL BIOTECHNOLOGY INC.  
XX (REUBER) REUBER L.  
XX (CREE) CREELMAN R.  
XX (PILG) PILGRIM M.  
XX (RIEC) RIECHMANN J L.  
XX (JIANG) JIANG C.  
XX (YUGG) YU G.  
XX (PINE) PINEDA O.  
XX (HEAR) HEARD J.  
XX  
XX Reuber L, Creelman R, Pilgrim M, Riechmann JL, Jiang C, Yu G;  
XX Pineda O, Heard J;  
XX WPI; 2001-335979/35.  
XX P-PSDB; AAE01932.  
XX  
XX Nucleic acids encoding plant transcription factor polypeptides, useful  
XX for altering the developmental and structural characteristics of  
XX plants, e.g. corn, potato and cotton plants -  
XX  
XX Claim 4; Page 113; 133pp; English.  
XX  
XX The present cDNA sequence is homologue G2422 of Arabidopsis thaliana  
XX transcription factor 6663. The transcription factors may be used to  
XX modify traits associated with structural or developmental  
XX characteristics of plants, e.g., corn, potato and cotton plants, when  
XX their expression level is altered. Specifically, they are used for  
XX modifying the nutritional content of plants or seeds, e.g., to modify  
XX vitamin and mineral content, to modify the oil content of seeds, to  
XX modify the insoluble sugar content of seeds, to modify phenyl lipid  
XX content, to modify fatty acid and modify wax content. They may also be  
XX used to alter seed characteristics such as shelf-life, size, stress  
XX tolerance, seedling vigour, pest and pathogen resistance and germination

CC rate. They may also be used in gene therapy. Therefore manipulating  
CC transcription factor levels in plants offers great potential in  
CC agricultural biotechnology for modifying a plant's traits.  
XX  
SQ Sequence 741 BP; 237 A; 122 C; 174 G; 208 T; 0 other;  
Query Match 22.5%; Score 185.4; DB 22; Length 741;  
Best Local Similarity 74.3%; Pred. No. 6.9e-36;  
Matches 234; Conservative 0; Mismatches 81; Indels 0; Gaps 0;  
QY 28 GGAGTGAGAAAAGTTCACTGACTGATGAAGAAATTTCTTTAAGAAAATGATGAT 87  
DB 19 GGGTTGAGAAAAGTCACTGACTGATGAAGAAATTTCTTTGAGGCAATGATGAT 78  
QY 88 AAGTATGAGAAAGAAATGGCATCTTTCCCAATGAGCTGCTGGAATGATGCGG 147  
DB 79 AAGTATGAGAAAGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCA 138  
QY 148 AAAAGTTGATGATGAGTGGCTGAATTTATCTAAGGCCCATATCAAGAGAGTGA 207  
DB 139 AAGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 198  
QY 208 GAAAGAGATGAGTGAATCTCATTTGAGGCTTCATPAGCTCTTAGGCAACAGATGCTCA 267  
DB 199 TGCTCCGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 258  
QY 268 CTATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327  
DB 259 TTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 318  
QY 328 AATCTTCTAAGAG 342  
DB 319 CATTTGATGAGAG 333

RESULT 12  
ABK65363  
ID ABK65363 standard; cDNA; 741 BP.  
XX  
XX AC ABK65363;  
XX  
XX 02-JUL-2002 (first entry)  
XX  
XX Arabidopsis cDNA encoding a transcription factor #215.  
XX  
XX DE Arabidopsis cDNA encoding a transcription factor #215.  
XX  
XX KW Plant; ss; gene; transcription factor; transgenic;  
XX agriculture; metabolic chemical; environmental stress; drought;  
XX microbial disease resistance; herbicide resistance; seed yield;  
XX fruit yield; growth rate; leaf senescence; flower senescence.  
XX  
XX OS Arabidopsis thaliana.  
XX  
XX PN WO200215675-A1.  
XX  
XX PD 28-FEB-2002.  
XX  
XX PF 22-AUG-2001; 2001MO-US26189.  
XX  
XX PR 22-AUG-2000; 2000US-227439P.  
XX 16-NOV-2000; 2000US-0713394.  
XX 16-APR-2001; 2001US-0837944.  
XX  
XX (MENDEL) MENDEL BIOTECHNOLOGY INC.  
XX (PILG) PILGRIM M.  
XX (CREE) CREELMAN R.  
XX (DUBE) DUBEL L A J.  
XX (HEAR) HEARD J.  
XX (JIANG) JIANG C.  
XX (KEDD) KEDDIE J.  
XX (ADAM) ADAM L.  
XX (RATC) RATCLIFF O.  
XX (REUB) REUBER J L.  
XX (RIEC) RIECHMANN J L.





PA (JIAN/) JIANG C.  
 PA (KEDD/) KEDDIE J.  
 PA (ADAM/) ADAM L.  
 PA (RATC/) RATCLIFF O.  
 PA (REUB/) REUBER J L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (YUGG/) YU G.  
 PA (PINE/) PINEDA O.  
 XX  
 PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J,  
 PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
 XX  
 DR WPI: 2002-292022/33.  
 DR P-PSDB: AAU93155.  
 XX  
 PT An isolated or recombinant polynucleotide used to produce a transgenic  
 PT plant -  
 XX  
 PS Claim 4; Page 792-794; 941pp; English.

CC The invention relates to 1 of 232 isolated or recombinant polynucleotides  
 CC encoding an Arabidopsis thaliana transcription factor, their variants,  
 CC complementary fragments, or related polynucleotide with 31% to 95%  
 CC sequence identity, where the plant possesses an altered trait as compared  
 CC to a wild-type or reference plant, or the plant exhibits an altered  
 CC phenotype as compared to a wild-type or reference plant, or the plant  
 CC exhibits ectopic expression or altered expression of one or more genes  
 CC associated with a plant trait as compared to a wild plant. Also included  
 CC are a transgenic plant comprising the polynucleotides, a computer  
 CC readable medium having stored sequence information, and identifying a  
 CC homologue sequence from a database comprising a plurality of known plant  
 CC sequences comprising inputting sequence information selected from one of  
 CC 464 fully defined sequences given in the specification. The isolated or  
 CC recombinant polynucleotide is used for producing a plant having a  
 CC modified trait, the method comprising selecting a polynucleotide that  
 CC encodes a polypeptide or an antisense nucleic acid, inserting the  
 CC polynucleotide or antisense nucleic acid into an expression vector,  
 CC introducing the vector into a plant or a cell of a plant to overexpress  
 CC the polypeptide or antisense nucleic acid, thereby producing a modified  
 CC plant, and selecting for a modified trait (e.g. increased  
 CC production of agriculturally useful proteins or metabolic chemicals,  
 CC pest tolerance, environmental stress response (e.g. drought), microbial  
 CC disease resistance, herbicide resistance, seed and fruit yield, growth  
 CC rate, leaf and flower senescence and many other traits listed in the  
 CC specification). The present sequence is one of the 232 polynucleotides  
 CC encoding an A. thaliana transcription factor.

XX  
 SQ Sequence 925 BP; 290 A; 160 C; 200 G; 275 T; 0 other;

Query Match 21.7%; Score 179; DB 24; Length 925;  
 Best Local Similarity 72.1%; Pred. No. 2.7e-34;  
 Matches 233; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 19 TCTTCATTGGAGTGAAGAAAGTTGATGACGTGATGAAGAAATTTCTTCAAGAAA 78  
 Db 68 TCGTCCAAAGGGCTGCGAAAAGTGTGACTGACCTAGAAAGATGCTCTTGAGACAG 127  
 QY 79 TGTATTGATTAAGTATGCTGAAGGAAATGCGATCTTGTCCCATTAAGAGCTGCTGAAT 138  
 Db 128 TGCATTATTAAGTATGCTGAAGGCAAAATGCGACCAAGTCTCTGTAAGAGCTGGGCTAAC 187  
 QY 139 AGATGTGCGAAAGTTGATGAGTGGTGGCTGAATTATCTAAGGCCCATATCAAGAGA 198  
 Db 188 CGGTGCGAGAAAGTTGATGAGTGGTGGCTGAATTATCTAAGGCCCATATCAAGAGA 247  
 QY 199 GGTGACTTTGAACAAGATGAAGTGAATCTCAATTTGAGGCTTCAAGCTCTTAGGCAAC 258  
 Db 248 GGAAGAACTTACCTGATGAAGTGCATCTTCTTGGCCTTCAAGGCTTCTAGGGAAT 307  
 QY 259 AGATGCTCACTTATTTGCTGAGACTTCCGGAAGAGACGTAACGATGGAATAACTAT 318  
 Db 308 AGTGGTCTTTAATTGCTGGAAGATTACTGTGCGACCGCAATGACGTCAGAAATTTAC 367

QY 319 TGAACACTAATCTTCTAAGGAA 341  
 Db 368 TGAACACTAATCTTCTAAGGAA 390

Search completed: January 30, 2004, 07:19:18  
 Job time: 243 secs



```

Db          422  TGAACACCATCT 435

RESULT 2
US-09-610-185C-1
; Sequence 1, Application US/09610185C
; Patent No. 6573432
; GENERAL INFORMATION:
; APPLICANT: Borevitz, Justin
; APPLICANT: Xia, Yiji
; APPLICANT: Dixon, Richard A.
; APPLICANT: Lamb, Christopher J.
; TITLE OF INVENTION: REGULATION OF ANTHOCYANIN PIGMENT
; FILE REFERENCE: SALKINS.003C1
; CURRENT APPLICATION NUMBER: US/09/610,185C
; CURRENT FILING DATE: 2000-07-05
; PRIOR APPLICATION NUMBER: US 09/603,244
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 956
; TYPE: DNA
; ORGANISM: Arabidopsis Thaliana
US-09-610-185C-1

```

```

Query Match 251; 21.8%; Score 180.2; DB 4; Length 956;
Best Local Similarity 68.0%; Pred. No. 8,7e-38;
Matches 251; Conservative 0; Mismatches 118; Indels 0; Gaps 0

QY 19 TCTTCATTGGGAGTGAGAAAAGTTTCATGCACTGATGAGAAAGATTTTCTTTAAAGAAA 78
Db 82 TCGTCCAAAAGGGCTCGAAAAGGTCCTTGACCTGACTGAAGAAAGATAGTCTTTGAGACAG 141
QY 79 TGTATTGATAGTATGAGGAAAGAAAATGCGACTCTTGTTCCCATAGAGCTGCTGAT 138
Db 142 TGCATTAATTAAGTATGAGAAAGGCAAAATGGCACCAAGTTCTGTAAAGCTGGGCTAAAC 201
QY 139 AGATGTCGGAAGAAAGTTAGATGATTGAGTGGCTGATTAATCTAAGGCCATATCAAGAGA 198
Db 202 CGGTCCAGGAAAAGTTGATAGTAAATGATGTTGAACATTTTGAACCAAGTATCAAGAGA 261
QY 199 GGTGACTTTGAAACAAGATGAAGTGATCTCATTTTGAGGCTTCATPAAGCTCTTAGGCAC 258
Db 262 GGAATAACTTAGCTCGATGAAGTCGATCTTCTTCCGCTTCATAGAGCTTCTAAGGAAT 321
QY 259 AGATGCTCACTTATGCTGCTGTAAGCTTCCCGAAGACACGCTAAAGATGTGAAAACTAT 318
Db 322 AGGTGGCTTTTAATGCTGGAAGATTAACCTGGTCGACCGCAAAAGACGTCAAGAAATTAC 381
QY 319 TGAACACTAATCTCTTAAGAAAGTTAAATACTACTAAATAATGTTCTCGCAAAAGATT 378
Db 382 TGGAAACACTCATCTGATGATGAGAAACATGAACCGTGTGTAAGATTAAGTGAAGAAAAGTA 441
QY 379 AACATAAG 387
Db 442 GACATTACG 450

RESULT 3
US-08-722-626B-1
: Sequence 1, Application US/08722626B
: Patent No. 5939601
: GENERAL INFORMATION:
: APPLICANT: Yang, Yining
: APPLICANT: Kleesig, Daniel, F.
: TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ENHANCED
: TITLE OF INVENTION: DISEASE RESISTANCE IN PLANTS
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dann, Dorfman, Herrell and Skillman
: STREET: 1601 Market Street Suite 720

```

CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2307  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTED Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/722,626E  
FILING DATE: 27-SEP-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pat Haean  
REGISTRATION NUMBER: 27,643  
REFERENCE/DOCKET NUMBER: 97-0010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215 563-4100  
TELEFAX: 215 563-4044  
TELEX:

```

? INFORMATION FOR SEQ ID NO: 1:
?
? SEQUENCE CHARACTERISTICS:
?
?     LENGTH: 1344 base pairs
?     TYPE: nucleic acid
?     STRANDEDNESS: single
?     TOPOLOGY: linear
?
?     MOLECULE TYPE: cDNA
?
?     HYPOTHETICAL: NO
?
?     ANTI-SENSE: NO
?
?     FRAGMENT TYPE:
?
?     ORIGINAL SOURCE:
?
?     FEATURE:
?
?     NAME/KEY: Coding Sequence
?     LOCATION: 148...981
?     OTHER INFORMATION:
?
? US-08-7122-626E-1

```

	Query Match	15.3%	Score 126.6	DB 2	Length 1344
	Best Local Similarity	59.2%	Pred. No. 8.5e-24		
	Matches 216	Conservative 0	Mismatches 149	Indels 0	Gaps 0
QY	24	ATTGGAGTGAAGAAAAGTTTCATGACTGATGAAGAAGATTTCTTCTTAAGAAAATGTAT	83		
Db	174	AATGGGCTGTAAGAAAAGGGCCATGATTCCTGGAAGAAGATCAGATTTCTCATCTCTTTCAT	233		
QY	84	TGATTAATATGATGTAAGAAAATTGGCATCTTGTTCCCATTAAGACTGGTCTGAATPAGATG	143		
Db	234	TCAAACTTAATGAGCCATGGCAATGGCCAGCCCTTCCAAAAGGCTGGACCTAATTGAGATG	293		
QY	144	TCGGAAAAGTTGTAGATTGAGTGGCTGAATTAATCTAAGGCCACATATCAAGAGAGTGA	203		
Db	294	CGGAGAGGTTTGACAGACTGGCGGTGACGAATTAATTGGGACAGATATAAAGAGGGGAAA	353		
QY	204	CTTTGAAACAAGATGAATGATCTCATTTTGAAGCTTCATAAGCTCTTAGGCACAGATG	263		
Db	354	TTTACCAAGAGMAAAMAGAAACAATTATCAGATTACATGAATAATGCTTGGCAATGATG	413		
QY	264	GTCATTAATTGCTGGTAGACTTCCCGAGAGACAGCTAACGATGTGAATAAATAATTGGAA	323		
Db	414	GTCGCAATAGCAGCAAAATTACAGAGACGAACAGACATGAATAATATAAATAATGTTGGCA	473		
QY	324	CACTAATCTTGAAGAAATTAAATCTACTAAATATGTTCTTCGCGAAAAGATTACAA	383		
Db	474	CACCCACTTGAAGAAAGACTCAAAAGATTATTAAGCTCTCTCAGAACTCCAAAAGACACTC	533		
QY	384	TAACT 388			
Db	534	CAAGT 538			



```
RESULT 4
US-09-008-979A-2
; Sequence 2, Application US/09008979A
; Patent No. 6080914
; GENERAL INFORMATION:
; APPLICANT: Comer, Timothy W.
; TITLE OF INVENTION: Strawberry Promoters and Genes
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carmen Rodriguez, Paralegal, Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/008,979A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kammerer, Patricia A.
; REGISTRATION NUMBER: 29,775
; REFERENCE/DOCKET NUMBER: MOBT:058 (38-21(10613)A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 787-1438
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-008-979A-2

Query Match      10.3%; Score 85.2; DB 3; Length 469;
Best Local Similarity 74.0%; Pred. No. 3.4e-13;
Matches 108; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 122 TAAGACTGCTCGATAGATGTCGAAAGTTGTAGATGAGTGGCTGAATTATCTAA 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 TACTTGACAGCTTAAACAGATGACGAGAGAGCTGTAGACTGAGATGGCTGAATTATTTAA 305

QY 182 GGCCACATATCAAGAGAGGTGACTTTGAACAAGATGAAGTATCTCATTTTGAGGCTTC 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 AGCCGAATATCAAGAGAGAGAGTTCGCGAGAGATGAAGTTCAATTGATCATCAGGCTTC 365

QY 242 ATAAAGCTTTAGGCAACAGATGCTCA 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 ATAAAGCTTTAGGCAACAGATTA 391

RESULT 5
US-09-460-618-2
; Sequence 2, Application US/09460618
; Patent No. 6235482
; GENERAL INFORMATION:
; APPLICANT: Comer, Timothy W.
; TITLE OF INVENTION: Strawberry Promoters and Genes
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carmen Rodriguez, Paralegal, Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/460,618
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US/09/008,979
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kammerer, Patricia A.
; REGISTRATION NUMBER: 29,775
; REFERENCE/DOCKET NUMBER: MOBT:058 (38-21(10613)A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 787-1438
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-460-618-2

Query Match      10.3%; Score 85.2; DB 3; Length 469;
Best Local Similarity 74.0%; Pred. No. 3.4e-13;
Matches 108; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 122 TAAGACTGCTCGATAGATGTCGAAAGTTGTAGATGAGTGGCTGAATTATCTAA 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 TACTTGACAGCTTAAACAGATGACGAGAGAGCTGTAGACTGAGATGGCTGAATTATTTAA 305

QY 182 GGCCACATATCAAGAGAGGTGACTTTGAACAAGATGAAGTATCTCATTTTGAGGCTTC 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 AGCCGAATATCAAGAGAGAGAGTTCGCGAGAGATGAAGTTCAATTGATCATCAGGCTTC 365

QY 242 ATAAAGCTTTAGGCAACAGATGCTCA 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 ATAAAGCTTTAGGCAACAGTAA 391

RESULT 6
US-09-310-235B-2
; Sequence 2, Application US/09310235B
; Patent No. 6392030
; GENERAL INFORMATION:
; APPLICANT: Comer, Timothy
; TITLE OF INVENTION: STRAWBERRY PROMOTERS AND GENES
; FILE REFERENCE: MOBT058-1 (11899,0058.DVUS00)
; CURRENT APPLICATION NUMBER: US/09/310,235B
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 469
; TYPE: DNA
; ORGANISM: fragaria x ananassa
;
US-09-310-235B-2

Query Match      10.3%; Score 85.2; DB 4; Length 469;
Best Local Similarity 74.0%; Pred. No. 3.4e-13;
Matches 108; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 122 TAAGACTGCTCGATAGATGTCGAAAGTTGTAGATGAGTGGCTGAATTATCTAA 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 TACTTGACAGCTTAAACAGATGACGAGAGAGCTGTAGACTGAGATGGCTGAATTATTTAA 305

QY 182 GGCCACATATCAAGAGAGGTGACTTTGAACAAGATGAAGTATCTCATTTTGAGGCTTC 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 AGCCGAATATCAAGAGAGAGAGTTCGCGAGAGATGAAGTTCAATTGATCATCAGGCTTC 365
```

QY 242 ATAGCTCTAGGACAGATGCTCA 267  
DB 366 ATAACTTCTAGGACAGATGCTCA 391

## RESULT 7

US-09-402-929-5  
; Sequence 5, Application US/09402929  
; Patent No. 6410825  
; GENERAL INFORMATION:  
; APPLICANT: Temple University - Of The Commonwealth System of Higher Education  
; APPLICANT: Toscani, Antonio  
; APPLICANT: Hatton, Kimi  
; APPLICANT: Reddy, E. P.  
; TITLE OF INVENTION: A-MYB NULL MUTANT TRANSGENIC ANIMALS AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
; STREET: Suite 1800 Two Penn Center Plaza  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/402,929  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US98/06896  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 6056-214 PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4880 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-402-929-5

Query Match 10.1%; Score 83.6; DB 4; Length 4880;

Best Local Similarity 56.4%; Pred. No. 2e-12;  
Matches 177; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

QY 28 GAGATGGAAGAAAGTTCATGACGATGAGAGATTTCTTCAAGAAATGATGAT 87  
DB 354 GAATGATTAAGGGCTTGACATTAAGAGAGATGAGAGGTTATTGATTTGTTCA 413  
QY 88 AAGTATGTAAGGAAATGATCTTGTCCATTAAGAGCTGCTGATAGATGTCG 147  
DB 414 AAATATGGCCAAAGATGCTTTTAATTC---AAACATTTAAAGAAATAGGC 470  
QY 148 AAAAGTTGATGAGTGTGCTGATTAATTAAGGCCAATATCAAGAGAGTGA 207  
DB 471 AAGCAGGTAGAGAAAGATGCAATATCATCTGATCTGAGTAAAGAAATCTTCTCG 530  
QY 208 GAAGAAGTGAAGTGAATCTGATTTTGAAGCTTCTAGGAAAGATGATGCA 267  
DB 531 ACAAGAGGAGGACAGATCATATGAACACATTAAGCGTTGGAAATCGTTGGCA 590  
QY 268 CTTATGCTGATGCTCCCGAAGACAGCTAAGATGGAATACTATTGGAACACT 327

DB 591 GAATTCGCAAACTACTTCAGAGAGACTGATTAATCTATCAAAAATCGAATTCT 650  
QY 328 AATCTCTAGGAA 341  
DB 651 ACTATGCAAGAAA 664

## RESULT 8

US-08-997-251-1  
; Sequence 1, Application US/08997251  
; Patent No. 6271440  
; GENERAL INFORMATION:  
; APPLICANT: GUBLER, FRANZ J.  
; APPLICANT: JACOBSEN, JOHN V.  
; TITLE OF INVENTION: PLANT REGULATORY PROTEINS III  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Boulevard  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/997,251  
; FILING DATE: 23-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO AU96/00383  
; FILING DATE: 21-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PN6470/95  
; FILING DATE: 09-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PN3779/95  
; FILING DATE: 23-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winer, Ellen P.  
; REGISTRATION NUMBER: 28,547  
; REFERENCE/DOCKET NUMBER: 110-97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2220 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Hordeum vulgare  
; STRAIN: Himalaya  
; DEVELOPMENTAL STAGE: Seed  
; TISSUE TYPE: Aleurone layer  
; IMMEDIATE SOURCE:  
; CLONE: HvGAMYb  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 275..1933  
; US-08-997-251-1

Query Match 9.5%; Score 78.4; DB 3; Length 2220;

Best Local Similarity 54.1%; Pred. No. 3.4e-11;  
Matches 160; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 32 TGAGAAAGCTTCATGACGATGAGAGATTTCTTCAAGAAATGATGATGAT 91

Db 399 TGAAGAGGCGCCCTGAGCTCCGCGAGAGAGCCCATCTGTGAGCTAGCTGAGAGAGC 458  
Qy 92 ATGTGAGAGAGAAATGCGATCTTGTCCATTAAGAGCTGTGTAATGATGCGAAAA 151  
Db 459 ACCGCGAGGGAAGTGAACCGCGTGCAGAGAAACACCGGCTGTCCGCGGCGAGA 518  
Qy 152 GTTGTAGATTGAGGTGCTGAATTATTAAGCCACATATCAAGAGAGTGTGACTTGAAC 211  
Db 519 GTGCGCGCTCCGCTGGCGAGACCACTCAGAGCCCACTCAAGAGAGGCGGCTTCAACC 578  
Qy 212 AAGATGAAGTGAATCTTATTTTGAAGCTTCAATGCTTTAGGCAACAGATGCTCACTTA 271  
Db 579 CCGAGGAGAGAGAGCTCATATCATCACTCACTCAAGATGGGCAACAGTGGGCTGGA 638  
Qy 272 TTGCTGTAGACTTCCCGGAGAGAGCAGCTAAGCTGTAAGAAATATTTGGAACACT 327  
Db 639 TGGCCGCTCATTTTCCAGGCGGTACTATTAATGAATTAAGAAATTTACTGGAACACT 694

## RESULT 9

US-09-402-929-1  
; Sequence 1, Application US/09402929  
; Patent No. 6410825  
; GENERAL INFORMATION:  
; APPLICANT: Temple University - Of The Commonwealth System of Higher Education  
; APPLICANT: Toscani, Antonio  
; APPLICANT: Hatton, Kimi  
; APPLICANT: Reddy, E. P.  
; TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
; STREET: Suite 1800 Two Penn Center Plaza  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/402,929  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US98/06896  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 6056-214 PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3602 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-402-929-1

Query Match 9.4%; Score 77.2; DB 4; Length 3602;  
Best Local Similarity 55.1%; Pred. No. 8.3e-11;  
Matches 173; Conservative 0; Mismatches 138; Indels 3; Gaps 1;  
Qy 28 GAGGTGAGAAAGGTTCAATGAGCTGATGAAGAAATTTTCTTCTTAAGAAATGATTTGAT 87  
Db 506 GAATGTATTAAGGGTCTTGTGACTTAAGAGAAAGATCATGAGAGGTTATTTGAATTTGATCAG 565

Qy 88 AACTATGTCAGAGAAATGCGATCTTGTCCATTAAGAGCTGTGTAATGATGTGG 147  
Db 566 AATATGGGCGAAAGAGGTGTCTTAATTTGC---AAACATTTAAAGAGAAATGAGC 622  
Qy 148 AAAAGTTGATGATTGAGGTGCTGAATTATCTTAAGCCACATATCAAGAGAGTGAATT 207  
Db 623 AACAGTGCAGAGAAAGATGCGCAATCACTGAACCTGGAAGTGAAGAGTCTTCTGG 682  
Qy 208 GAACAGATGAAGTGAATCTATTTGAGGCTTCATTAAGCTCTTAGGCAACAGATGCTCA 267  
Db 683 ACGAAGAGAGAGACAGATCATATTAATGAAGACACAGAGGCTTGAGAAACGTTGGCC 742  
Qy 268 CTATTTGCTGTAGACTTCCCGAGAGAGAGCTAAGATGTGAAGAAATTAATTTGAACT 327  
Db 743 GAATTTGCTTAAGTTACTTCTTGGAGAGAGCTGATTAATTTATCAAAATCATTTGAAATTC 802  
Qy 328 AATCTTGAAGAA 341  
Db 803 ACCATGCGAAGAA 816

## RESULT 10

US-09-402-929-4  
; Sequence 4, Application US/09402929  
; Patent No. 6410825  
; GENERAL INFORMATION:  
; APPLICANT: Temple University - Of The Commonwealth System of Higher Education  
; APPLICANT: Toscani, Antonio  
; APPLICANT: Hatton, Kimi  
; APPLICANT: Reddy, E. P.  
; TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
; STREET: Suite 1800 Two Penn Center Plaza  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/402,929  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US98/06896  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 6056-214 PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6775 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-402-929-4

Query Match 9.4%; Score 77.2; DB 4; Length 6775;  
Best Local Similarity 55.1%; Pred. No. 1e-10;  
Matches 173; Conservative 0; Mismatches 138; Indels 3; Gaps 1;  
Qy 28 GAGGTGAGAAAGGTTCAATGAGCTGATGAAGAAATTTTCTTCTTAAGAAATGATTTGAT 87

Db 3755 GAATGTGATAAGGCTCTTGACTAAGAGAGATGAGGGTTATGTAATTAGTTCAAG 3814  
Qy 88 AAGTATGATGAAGAAATGAGCATCTTGTCCATAGAGCTGCTGAATAGATGTCGG 147  
Db 3815 AATATATGGCCAAAAGATGCTCTTAAATTCG---AAAACCTTTAAAGAGAAATAGCC 3871  
Qy 148 AAAAGTTGATGATGAGCTGCTGAATTAATCTAAGGCCACATATCAAGAGAGTGAATTT 207  
Db 3872 AAGAGTGCAGAGAAATGAGCAATACCTGAACCTGAAAGTGAAGAGTCTTCTCG 3931  
Qy 208 GAACAATGAATGATGATCTATTTGAGCTTCATAGCTCTTAAAGCAACATGCTCA 267  
Db 3932 ACAGAAAGAAAGACAGATCTATATGAGCAACAGAGCCCTGGAAACCGTTGGGCC 3991  
Qy 268 CTTATGCTGATGATCTTCCGGAAGACAGCTAAGATGAAATACTATTGGACACT 327  
Db 3992 GAGATTGCTAAGTACTTCTCGAAGAGATGATATTTCTATCAAAAATCATTTGAAATCT 4051  
Qy 328 AATCTCTAAGAA 341  
Db 4052 ACATGCGAAGAAA 4065

RESULT 11  
US-08-997-251-3  
Sequence 3, Application US/08997251  
Patent No. 6271440  
GENERAL INFORMATION:  
APPLICANT: GUEBLER, FRANZ J.  
APPLICANT: JACOBSEN, JOHN V.  
TITLE OF INVENTION: PLANT REGULATORY PROTEINS III  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Boulevard  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/997,251  
FILING DATE: 23-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO AU96/00383  
FILING DATE: 21-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PN6470/95  
FILING DATE: 09-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PN3779/95  
FILING DATE: 23-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Winner, Ellen P.  
REGISTRATION NUMBER: 28,547  
REFERENCE/DOCKET NUMBER: 110-97  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2352 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Oryza sativa

FEATURE:  
NAME/KEY: CDS  
LOCATION: 396..2054  
US-08-997-251-3  
Query Match 9.1%; Score 75.2; DB 3; Length 2352;  
Best Local Similarity 53.4%; Pred. No. 2,4e-10;  
Matches 158; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

Qy 32 TGAGAAAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 91  
Db 514 TGAAGAGGGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 573  
Qy 92 ATGTGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 151  
Db 574 ACGGAGGGGAACTGAAAGCGGTGAGAAAGAACCCGGCTGTCGGTGGCGAAG 633  
Qy 152 GTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 211  
Db 634 GCTGCCGCTCCGTTGGGCGAACCTGAGGCCAACCTCAAGAAAGGGGCTTCAACG 693  
Qy 212 AAGATGAAGTGAATCTATTTGAGGCTTCATAGCTCTTAAAGCAACAGATGCTCACTTA 271  
Db 694 CCGAGAGAGAGAGCTCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 753  
Qy 272 TTGCTGTAGACTTCCGGAAGACAGCTAAGATGTAATACTATTGGACACT 327  
Db 754 TGCGCGCTCATTTGCCAGGGCGCATGATATGAATTAAGAAATTAAGTGAATTAAT 809

RESULT 12  
US-09-008-979A-6  
Sequence 6, Application US/09008979A  
Patent No. 6080914  
GENERAL INFORMATION:  
APPLICANT: Comer, Timothy W.  
TITLE OF INVENTION: Strawberry Promoters and Genes  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carmen Rodriguez, Paralegal, Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/008,979A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Kammerer, Patricia A.  
REGISTRATION NUMBER: 29,775  
REFERENCE/DOCKET NUMBER: MOBT:058 (38-21(10613)A)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (713) 787-1438  
TELEFAX: (713) 789-2679  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 785 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-008-979A-6

Query Match 8.7%; Score 71.4; DB 3; Length 785;  
Best Local Similarity 61.3%; Pred. No. 1.6e-09;  
Matches 133; Conservative 0; Mismatches 81; Indels 3; Gaps 1;

QY 220 GIGGATCTCATTTTGGAGCTTTCATTAAGCTCTTAAGCAAGATGTCACCTTAATGCTGT 279  
| | | | |  
DB 2 GTTGATTTGATCATCGAGCTTTCATTAAGCTCTTAAGCAAGATGTCACCTTAATGCTGT 61  
| | | | |  
QY 280 AGACTTCCCGAAGAGCAAGCTTAAGCAAGATGTCACCTTAATGCTGT 339  
| | | | |  
DB 62 CGATTCGAGAGAGCAAGCTTAAGCAAGATGTCACCTTAATGCTGT 118  
| | | | |  
QY 340 AAGTAAATCTACTAAATTTGTTCTCGGAAAGATTAACATAAGTGTGGAAT 399  
| | | | |  
DB 119 AAAAGATCAAAAGAGCGCTTTCATTAAGCAAGATGTCACCTTAATGCTGT 178  
| | | | |  
QY 400 AGTACTAAGTGAATTAATTAAGCAAGCA 436  
| | | | |  
DB 179 ACAATAGCTTACCAATTTGAAGCTCGACCAAGCA 215  
| | | | |

RESULT 13  
US-09-460-618-6  
; Sequence 6, Application US/09460618  
; Patent No. 6235482

GENERAL INFORMATION:

APPLICANT: Conner, Timothy W.

TITLE OF INVENTION: Strawberry Promoters and Genes

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carmen Rodriguez, Paralegal, Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/460,618

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/008,979

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Kammerer, Patricia A.

REGISTRATION NUMBER: 29,775

REFERENCE/DOCKET NUMBER: MOBT:058 (38-21 (10613)A)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (713) 787-1438

TELEFAX: (713) 789-2679

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 785 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-09-460-618-6

Query Match 8.7%; Score 71.4; DB 3; Length 785;  
Best Local Similarity 61.3%; Pred. No. 1.6e-09;  
Matches 133; Conservative 0; Mismatches 81; Indels 3; Gaps 1;

QY 220 GIGGATCTCATTTTGGAGCTTTCATTAAGCTCTTAAGCAAGATGTCACCTTAATGCTGT 279  
| | | | |  
DB 2 GTTGATTTGATCATCGAGCTTTCATTAAGCTCTTAAGCAAGATGTCACCTTAATGCTGT 61  
| | | | |  
QY 280 AGACTTCCCGAAGAGCAAGCTTAAGCAAGATGTCACCTTAATGCTGT 339  
| | | | |  
DB 62 CGATTCGAGAGAGCAAGCTTAAGCAAGATGTCACCTTAATGCTGT 118  
| | | | |  
QY 340 AAGTAAATCTACTAAATTTGTTCTCGGAAAGATTAACATAAGTGTGGAAT 399  
| | | | |

DB 119 AAAAGATCAAAAGAGCGCTTTCATTAAGCAAGATGTCACCTTAATGCTGT 178  
| | | | |  
QY 400 AGTACTAAGTGAATTAATTAAGCAAGCA 436  
| | | | |  
DB 179 ACAATAGCTTACCAATTTGAAGCTCGACCAAGCA 215  
| | | | |

RESULT 14  
US-09-310-235B-6  
; Sequence 6, Application US/09310235B  
; Patent No. 6392030

GENERAL INFORMATION:

APPLICANT: Conner, Timothy

TITLE OF INVENTION: STRAWBERRY PROMOTERS AND GENES

FILE REFERENCE: MOBT058-1 (11899, 0058, DVS00)

CURRENT APPLICATION NUMBER: US/09/310,235B

CURRENT FILING DATE: 1998-05-12

NUMBER OF SEQ ID NOS: 14

SOFTWARE: Patent in version 3.0

SEQ ID NO 6

LENGTH: 785

TYPE: DNA

ORGANISM: fragaria x ananassa

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)-(785)

OTHER INFORMATION: N = A, T, C or G

US-09-310-235B-6

Query Match 8.7%; Score 71.4; DB 4; Length 785;  
Best Local Similarity 61.3%; Pred. No. 1.6e-09;  
Matches 133; Conservative 0; Mismatches 81; Indels 3; Gaps 1;

QY 220 GIGGATCTCATTTTGGAGCTTTCATTAAGCTCTTAAGCAAGATGTCACCTTAATGCTGT 279  
| | | | |  
DB 2 GTTGATTTGATCATCGAGCTTTCATTAAGCTCTTAAGCAAGATGTCACCTTAATGCTGT 61  
| | | | |  
QY 280 AGACTTCCCGAAGAGCAAGCTTAAGCAAGATGTCACCTTAATGCTGT 339  
| | | | |  
DB 62 CGATTCGAGAGAGCAAGCTTAAGCAAGATGTCACCTTAATGCTGT 118  
| | | | |  
QY 340 AAGTAAATCTACTAAATTTGTTCTCGGAAAGATTAACATAAGTGTGGAAT 399  
| | | | |  
DB 119 AAAAGATCAAAAGAGCGCTTTCATTAAGCAAGATGTCACCTTAATGCTGT 178  
| | | | |  
QY 400 AGTACTAAGTGAATTAATTAAGCAAGCA 436  
| | | | |  
DB 179 ACAATAGCTTACCAATTTGAAGCTCGACCAAGCA 215  
| | | | |

RESULT 15  
US-08-232-463-14/C  
; Sequence 14, Application US/08232463  
; Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.

APPLICANT: SCHREIFLINGER, F.

TITLE OF INVENTION: RECOMBINANT FOMLPX VIRUS

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

```
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZSPC-F1s
US-08-232-463-14

Query Match      8.4%; Score 69.6; DB 1; Length 7218;
Best Local Similarity 7.0%; Pred.No. 1e-08;
Matches 30; Conservative 233; Mismatches 167; Indels 0; Gaps 0;

QY 305 ATGTGAAAACCTATTGGAACACTAATCTTCTAAGAACTTAATACTACTAATGTTTC 364
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1497 ATGTAGGCATCTAGTATTAATCTATCTATGCAAGTAGTTAAGATAGAGATTGG 1438
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 365 CTGCGGAAAAGATTAACTAAGTGTGAGAAATTAGTACTAGATTGAATTATAAAC 424
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1437 TACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1378
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 425 CTGACGACGCAATATTTCACAGCACAATGAAAGATTGTTACAACAATATGTAATT 484
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1377 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1318
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 485 TGAAGAGAGAGCAATTGCAAGAAATAAAGTGAAGAAACAATCCAGATGCATGCA 544
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1317 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1258
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 545 TGAACAACGTAGATTCATGTGATAAATTACTGAAAAATGCAATGACGATATTGAAG 604
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1257 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1198
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 605 AAGATGAAGAGTGTATTAATATGAAAAAAACCTAACAACTTGTACATGAGAAA 664
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1197 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1138
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 665 TATCACCACTAATATTGTAAGTAATCCATGACCAAGCAAAATAAGTATGATG 724
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1137 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1078
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 725 AAAATTGGGG 734
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1077 RRRRRRRRRR 1068
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: January 30, 2004, 08:33:09  
Job time : 64 secs



D	b	181	AGGCCACATATCAAGAGAGGTGACTTGGACAAATGAAGTGATCTCATTTTGAAGCTT	240
Q	y	241	CATAGAGCTTTAGGCAACAGATGGTCACTTAATGCTGTAGACTTCCCGAAGGACGCT	300
D	b	241	CATAGAGCTTTAGGCAACAGATGGTCACTTAATGCTGTAGACTTCCCGAAGGACGCT	300
Q	y	301	AACGATGTGAAAACTATTGGACAACCTAATCTTCTAAGGAAGTTAATACTACTAAATT	360
D	b	301	AACGATGTGAAAACTATTGGACAACCTAATCTTCTAAGGAAGTTAATACTACTAAATT	360
Q	y	361	GTTCCTGCGCAAAAGATTAACTAATAGTGTGAGAAATTACTACTAAGATTGAATTATA	420
D	b	361	GTTCCTGCGCAAAAGATTAACTAATAGTGTGAGAAATTACTACTAAGATTGAATTATA	420
Q	y	421	AAACCTCAACGACCGCAAGTATTTCTCAAGCAACAATGAAGATGTTACAAACATTAATGTA	480
D	b	421	AAACCTCAACGACCGCAAGTATTTCTCAAGCAACAATGAAGATGTTACAAACATTAATGTA	480
Q	y	481	ATTTTGGACGAGAGAGAACTATTGCAAGGAAATTAATAGTGAAAACAAACCTCCAGATGCA	540
D	b	481	ATTTTGGACGAGAGAGAACTATTGCAAGGAAATTAATAGTGAAAACAAACCTCCAGATGCA	540
Q	y	541	TCCGATGCAACACGTAGATCCATGGTGTGATAAATTACTGGAAAATTGCAATGACGATAT	600
D	b	541	TCCGATGCAACACGTAGATCCATGGTGTGATAAATTACTGGAAAATTGCAATGACGATAT	600
Q	y	601	GAAAGAGTGAAGAGGTTGTAATTAAATTATGAAAAACACTAACAGTTTGTTCATGAA	660
D	b	601	GAAAGAGTGAAGAGGTTGTAATTAAATTATGAAAAACACTAACAGTTTGTTCATGAA	660
Q	y	661	GAAATATCACCACTTAATTAATTTGTGTGAAGTAACTCCATGCAACAAGACAATAAGT	720
D	b	661	GAAATATCACCACTTAATTAATTTGTGTGAAGTAACTCCATGCAACAAGACAATAAGT	720
Q	y	721	CATGAAATTTGGGTGAATTTTCTCTTAATTACCAACCATGCAACAAGAGTCAAAAT	780
D	b	721	CATGAAATTTGGGTGAATTTTCTCTTAATTACCAACCATGCAACAAGAGTCAAAAT	780
Q	y	781	GATGATTTTCTGCTGGAATTGACTTAATGGAATCTACTGTGATTA	825
D	b	781	GATGATTTTCTGCTGGAATTGACTTAATGGAATCTACTGTGATTA	825
RESULT 2				
US-10-033-190-4				
Sequence 4, Application US/10033190				
Publication No. US20020133848A1				
GENERAL INFORMATION:				
APPLICANT: Exelixis Plant Sciences, Inc.				
TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF AN ANTHOCYANIN MUTANT (ANTHOCYANIN)				
TITLE OF INVENTION: TOMATO				
FILE REFERENCE: EP01-002C				
CURRENT APPLICATION NUMBER: US/10/033.190				
PRIORITY FILING DATE: 2001-10-29				
PRIORITY FILING DATE: US 60/244,685				
PRIORITY FILING DATE: 2000-10-30				
NUMBER OF SEQ ID NOS: 5				
SOFTWARE: PatentIn version 3.1				
SEQ ID NO 4				
LENGTH: 1012				
TYPE: DNA				
ORGANISM: Lycopersicon esculentum				
US-10-033-190-4				

Oy	61	GATTTTCTTCAAGAAAATGATTGATAGATGGTGAAGGAAAATGGCATCTTGTCC	120
Db	61	GATTTTCTTCAAGAAAATGATTGATAGATGGTGAAGGAAAATGGCATCTTGTCC	120
Oy	121	ATAAGAGT-----	129
Db	121	ATAAGAGT-----	129
Oy	121	ATAAGAGT-----	129
Db	121	ATAAGAGT-----	129
Oy	130	-----GCTGAAATGATGT	144
Db	130	-----GCTGAAATGATGT	144
Oy	145	CGGAAAAGTTGATAGTGGGTGGTGAATATCTPAAGGCCACATPAAGAGAGTGC	204
Db	145	CGGAAAAGTTGATAGTGGGTGGTGAATATCTPAAGGCCACATPAAGAGAGTGC	204
Oy	205	TTTGAACAAGTGAAGTGAATCTCATTTTGGAGCTTCATPAAGCTTTAGGCAAC-----	258
Db	205	TTTGAACAAGTGAAGTGAATCTCATTTTGGAGCTTCATPAAGCTTTAGGCAAC-----	258
Oy	259	-----	258
Db	259	-----	258
Oy	361	GCAAGTTATGTTTGAACAATAATTGATTAGTATATATATATACGTGACTATTTTC	420
Db	361	GCAAGTTATGTTTGAACAATAATTGATTAGTATATATATATATACGTGACTATTTTC	420
Oy	259	-----AGATGCTACTTATTCGTGGTACCTCCCGGAAG	293
Db	259	-----AGATGCTACTTATTCGTGGTACCTCCCGGAAG	293
Oy	421	ATCTAAATGTTACGTTATTTTACGTAATGGTCACTTATTCGTGGTACCTCCCGGAAG	480
Db	421	ATCTAAATGTTACGTTATTTTACGTAATGGTCACTTATTCGTGGTACCTCCCGGAAG	480
Oy	294	GACAGCTPAAGATGTGAAAACTPATTTGGAACTCATATCTTCAAGGAAGTTAAATACAC	353
Db	294	GACAGCTPAAGATGTGAAAACTPATTTGGAACTCATATCTTCAAGGAAGTTAAATACAC	353
Oy	481	GACAGCTPAAGATGTGAAAACTPATTTGGAACTCATATCTTCAAGGAAGTTAAATACAC	540
Db	481	GACAGCTPAAGATGTGAAAACTPATTTGGAACTCATATCTTCAAGGAAGTTAAATACAC	540
Oy	354	TAAATTTGTTCTCGCGAAAAGATTAACAATPAAGTGGAGAAATTAGTACTPAAGTTGA	413
Db	354	TAAATTTGTTCTCGCGAAAAGATTAACAATPAAGTGGAGAAATTAGTACTPAAGTTGA	413
Oy	541	TAAATTTGTTCTCGCGAAAAGATTAACAATPAAGTGGAGAAATTAGTACTPAAGTTGA	600
Db	541	TAAATTTGTTCTCGCGAAAAGATTAACAATPAAGTGGAGAAATTAGTACTPAAGTTGA	600
Oy	414	AATTATATAAACTCAACGACCGCAATATTTCTCAAGCACAAATGAAGAATGTTACAAACA	473
Db	414	AATTATATAAACTCAACGACCGCAATATTTCTCAAGCACAAATGAAGAATGTTACAAACA	473
Oy	601	AATTATATAAACTCAACGACCGCAATATTTCTCAAGCACAAATGAAGAATGTTACAAACA	660
Db	601	AATTATATAAACTCAACGACCGCAATATTTCTCAAGCACAAATGAAGAATGTTACAAACA	660
Oy	474	TAAATTTGTTCTCGCGAAAAGATTAACAATPAAGTGGAGAAATTAGTACTPAAGTTGA	533
Db	474	TAAATTTGTTCTCGCGAAAAGATTAACAATPAAGTGGAGAAATTAGTACTPAAGTTGA	533
Oy	661	TAAATTTGTTCTCGCGAAAAGATTAACAATPAAGTGGAGAAATTAGTACTPAAGTTGA	720
Db	661	TAAATTTGTTCTCGCGAAAAGATTAACAATPAAGTGGAGAAATTAGTACTPAAGTTGA	720
Oy	534	AGATGCAATCGATGACAACTAGATCCATGGTGGATTAATTTACTCGAAAAATTGCATGA	593
Db	534	AGATGCAATCGATGACAACTAGATCCATGGTGGATTAATTTACTCGAAAAATTGCATGA	593
Oy	721	AGATGCAATCGATGACAACTAGATCCATGGTGGATTAATTTACTCGAAAAATTGCATGA	780
Db	721	AGATGCAATCGATGACAACTAGATCCATGGTGGATTAATTTACTCGAAAAATTGCATGA	780
Oy	594	CGATATTGAAGAAGATGAAGGTTGTAATTATATGAAAAAACACTPAACAAGTTGTT	653
Db	594	CGATATTGAAGAAGATGAAGGTTGTAATTATATGAAAAAACACTPAACAAGTTGTT	653
Oy	781	CGATATTGAAGAAGATGAAGGTTGTAATTATATGAAAAAACACTPAACAAGTTGTT	840
Db	781	CGATATTGAAGAAGATGAAGGTTGTAATTATATGAAAAAACACTPAACAAGTTGTT	840
Oy	654	ACATGAAGAATATATCACCACTTAAATATTTGGTGAAGTAACTCCATGCCAACAAAGACA	713
Db	654	ACATGAAGAATATATCACCACTTAAATATTTGGTGAAGTAACTCCATGCCAACAAAGACA	713
Oy	841	ACATGAAGAATATATCACCACTTAAATATTTGGTGAAGTAACTCCATGCCAACAAAGACA	900
Db	841	ACATGAAGAATATATCACCACTTAAATATTTGGTGAAGTAACTCCATGCCAACAAAGACA	900
Oy	714	AATTAAGTCATAAAATTTGGGGGGAATTTCTCTTAATTTTACCAACCATGCAACAAGAGCT	773
Db	714	AATTAAGTCATAAAATTTGGGGGGAATTTCTCTTAATTTTACCAACCATGCAACAAGAGCT	773
Oy	901	AATTAAGTCATAAAATTTGGGGGGAATTTCTCTTAATTTTACCAACCATGCAACAAGAGCT	960
Db	901	AATTAAGTCATAAAATTTGGGGGGAATTTCTCTTAATTTTACCAACCATGCAACAAGAGCT	960
Oy	774	ACAAAATGATATTTTCTGCTGAATTTGACTTATGAAATCTACTTGATTAA	825
Db	774	ACAAAATGATATTTTCTGCTGAATTTGACTTATGAAATCTACTTGATTAA	825
Oy	961	ACAAAATGATATTTTCTGCTGAATTTGACTTATGAAATCTACTTGATTAA	1012
Db	961	ACAAAATGATATTTTCTGCTGAATTTGACTTATGAAATCTACTTGATTAA	1012

	Query Match	Score	DB	Length
Best Local Similarity	74.9%	618	14	1012
Matches	825	Conservative	0	Mismatches 0; Indels 187; Gaps 2

```

RESULTS
US-09-934-455-479
; Sequence 479, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Duwell, Arnold

```



APPLICANT: Heard, Jacqueline  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Keddie, James  
APPLICANT: Pilgrim, Marsha  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Reuber, Lynne  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Yu, Guo-Liang  
APPLICANT: Pineda, Omeira  
TITLE OF INVENTION: Genes for Modifying Plant Traits IV  
FILE REFERENCE: MBI-0025  
CURRENT APPLICATION NUMBER: US/09/934,455  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/227439  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: MBI-0022  
PRIOR FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: MBI-0023  
PRIOR FILING DATE: 2001-04-17  
NUMBER OF SEQ ID NOS: 516  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO: 479  
LENGTH: 762  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(630)  
OTHER INFORMATION: G2421  
US-09-934-455-479

Query Match 23.0%; Score 189.6; DB 11; Length 762;  
Best Local Similarity 74.1%; Pred. No. 1.6e-33;  
Matches 240; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

19 TCTTCATTGGAGTGAAGAAAGGTTTCAGTATGAGAGATTTCTTCTAAGAAA 78  
10 TCGTCCAAAGGGTTGAGAAAGTGATGATGCTGGAAGAAATAGTCTTTAGAGCAG 69  
79 TGTATTGATTAAGTATGCTGAGAGAAATGCGATCTTTGCCATAAGAGCTGCTGAAT 138  
70 TGTATTGATTAAGTATGAGAGAAATGCGATCTTTGCCATAAGAGCTGCTGAAT 129  
139 AGATGTCGAGAAAGTTGATGATGAGTGGCTGAATTTATCTAAGGCACATCAAGAG 198  
130 CGGTGAGAGAAAGTTGATGATGAGTGGCTGAATTTATCTAAGGCACATCAAGAG 189  
199 GGTGACTTTGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 258  
190 GGAATTTTAAAGTCTGATGAGTGAATCTTCTTCTGCTTCTAAGAGTCTTGAAGAA 249  
259 AGATGTCATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 318  
250 AGGTGTCCTTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 309  
319 TGAACACTTAATCTTCTAAGAGAG 342  
310 TGAACACCCATCTGATGAGAGAG 333

RESULT 4  
US-09-770-445-268/C  
Sequence 268, Application US/09770445  
Patent No. US20020023281A1  
GENERAL INFORMATION:  
APPLICANT: Gorlach, Jorn  
APPLICANT: An, Yong-Qiang  
APPLICANT: Hamilton, Carol M.  
APPLICANT: Price, Jennifer L.  
APPLICANT: Raines, Tracy M.  
APPLICANT: Yu, Yang  
APPLICANT: Rameaka, Joshua G.  
APPLICANT: Page, Amy

APPLICANT: Matthew, Abraham V.  
APPLICANT: Ledford, Brooke L.  
APPLICANT: Woessner, Jeffrey P.  
APPLICANT: Haas, William David  
APPLICANT: Garcia, Carlos A.  
APPLICANT: Kricker, Maja  
APPLICANT: Slader, Ted  
APPLICANT: Davis, Keith R.  
APPLICANT: Allen, Keith  
APPLICANT: Hoffman, Neil  
APPLICANT: Horden, Patrick  
TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
FILE REFERENCE: 2023US (PAPA-012PRV)  
CURRENT APPLICATION NUMBER: US/09/770,445  
CURRENT FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: US 60/178,472  
PRIOR FILING DATE: 2000-01-27  
NUMBER OF SEQ ID NOS: 999  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 268  
LENGTH: 982  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(982)  
OTHER INFORMATION: n = A,T,C or G  
US-09-770-445-268

Query Match 22.5%; Score 186; DB 9; Length 982;  
Best Local Similarity 74.5%; Pred. No. 1.2e-32;  
Matches 234; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

19 TCTTCATTGGAGTGAAGAAAGGTTTCAGTATGAGAGATTTCTTCTAAGAAA 78  
911 TCGTCCAAAGGGTTGAGAAAGTGATGATGATGATGATGATGATGATGATGATG 852  
79 TGTATTGATTAAGTATGCTGAGAGAAATGCGATCTTTGCCATAAGAGCTGCTGAAT 138  
851 TGTATTGATTAAGTATGAGAGAAATGCGATCTTTGCCATAAGAGCTGCTGAAT 792  
139 AGATGTCGAGAAAGTTGATGATGAGTGGCTGAATTTATCTAAGGCACATCAAGAG 198  
791 CGATGAGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 732  
199 GGTGACTTTGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 258  
731 GGAATTTTAAAGTCTGATGAGTGAATCTTCTTCTGCTTCTAAGAGTCTTGAAGAA 672  
259 AGATGTCATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 318  
671 AGGTGTCCTTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 612  
319 TGAACACTTAATCT 332  
611 TGAACACCCATCT 598

RESULT 5  
US-10-295-403-43  
Sequence 43, Application US/10295403  
Publication No. US20030101481A1  
GENERAL INFORMATION:  
APPLICANT: Heard, Jacqueline  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Adam, Luc  
APPLICANT: Broun, Pierre  
APPLICANT: Reuber, Lynne  
APPLICANT: Pineda, Omeira  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Keddie, James  
APPLICANT: Zhang, James

```
APPLICANT: Benito, Maria-Ines
APPLICANT: Yu, Guo-Liang
APPLICANT: Fromm, Mike
TITLE OF INVENTION: PLANT GENE SEQUENCES I
FILE REFERENCE: MBI-0003
CURRENT APPLICATION NUMBER: US/10/295,403
CURRENT FILING DATE: 2002-11-15
PRIOR APPLICATION NUMBER: US/09/394,519
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: 60/101,349
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: 60/103,312
PRIOR FILING DATE: 1998-10-06
PRIOR APPLICATION NUMBER: 60/108,734
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/113,409
PRIOR FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 43
LENGTH: 1016
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: (96)..(842)
OTHER INFORMATION: G663
US-10-295-403-43
```

```
Query Match      22.5%; Score 186; DB 15; Length 1016;
Best Local Similarity 74.5%; Pred. No. 1.2e-32;
Matches 234; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
```

```
QY 19 TCTTCATTGGAGTGAAGAAAGGTTCTATGACATGATGAAGAATTTCTTCTTAAGAAA 78
DB 105 TCTGCCAAAGGGTTGAGAAAGGTCATGACCTGCTAAGAAAGATCTCTTGAAGGCTA 164
QY 79 TGTATTGATAGTATGTTGGAAGAAATGTCATCTTCTCCATTAAGAGCTGCTGAAT 138
DB 165 TGTATTGATAGTATGGAAGAAAGGCAATGGCATCAAGTTCTTGAAGAGCTGGCTAAT 224
QY 139 AGATGTCGAAAGTTGTAGATTGAGGTGCTGAATTAATCAAGCCACATATCAAGAGA 198
DB 225 CGATGCAAGAAAGGTTGTAGCTTAAGATGTTGAACATATTGAAGCCAAAGATCAAGAGA 284
QY 199 GGTGACTTTGAACAAGATGAAGTGAATCTCATTTTGAAGCTTCAATAGCTTTAGGCAAC 258
DB 285 GGAAGACTTAGCATATGATGAAGTTGATCTTCTTGCCCTTCAATAGCTTTAGGAAAT 344
QY 259 AGATGTCACCTTATTGCTGGTAGACTCCCGGAAGGACAGTAACATGTAAGAAACTAT 318
DB 345 AGGTGGTCTCTTGAATGCTGGTGCATGCTGGTGGACCGCTAATGATGTCAAAATTAC 404
QY 319 TGGAACTAATCT 332
DB 405 TGGAACTAATCT 418
```

```
RESULT 6
US-09-934-455-13
; Sequence 13, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
```

```
APPLICANT: Yu, Guo-Liang
APPLICANT: Pineda, Omaira
TITLE OF INVENTION: Genes for Modifying Plant Traits IV
FILE REFERENCE: MBI-0025
CURRENT APPLICATION NUMBER: US/09/934,455
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/227439
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: MBI-0022
PRIOR FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: MBI-0023
PRIOR FILING DATE: 2001-04-17
NUMBER OF SEQ ID NOS: 516
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 1033
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: (113)..(862)
OTHER INFORMATION: G663
US-09-934-455-13
```

```
Query Match      22.5%; Score 186; DB 11; Length 1033;
Best Local Similarity 74.5%; Pred. No. 1.2e-32;
Matches 234; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
```

```
QY 19 TCTTCATTGGAGTGAAGAAAGGTTCTATGACATGATGAAGAATTTCTTCTTAAGAAA 78
DB 122 TCTGCCAAAGGGTTGAGAAAGGTCATGACCTGCTAAGAAAGATGCTCTTGAAGGCTA 181
QY 79 TGTATTGATAGTATGTTGGAAGAAATGTCATCTTCTCCATTAAGAGCTGCTGAAT 138
DB 182 TGTATTGATAGTATGGAAGAAAGGCAATGGCATCAAGTTCTTGAAGAGCTGGCTAAT 241
QY 139 AGATGTCGAAAGTTGTAGATTGAGGTGCTGAATTAATCAAGCCACATATCAAGAGA 198
DB 242 CGATGCAAGAAAGGTTGTAGCTTAAGATGTTGAACATATTGAAGCCAAAGTATCAAGAGA 301
QY 199 GGTGACTTTGAACAAGATGAAGTGAATCTCATTTTGAAGCTTCAATAGCTTTAGGCAAC 258
DB 302 GGAAGACTTAGCATATGATGAAGTTGATCTTCTTGCCCTTCAATAGCTTTAGGAAAT 361
QY 259 AGATGTCACCTTATTGCTGGTAGACTCCCGGAAGGACAGTAACATGTAAGAAACTAT 318
DB 362 AGGTGGTCTCTTGAATGCTGGTGCATGCTGGTGGACCGCTAATGATGTCAAAATTAC 421
QY 319 TGGAACTAATCT 332
DB 422 TGGAACTAATCT 435
```

```
RESULT 7
US-09-934-455-481
; Sequence 481, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omaira
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
```

```

?
? CURRENT FILING DATE: 2001-08-22
? PRIOR APPLICATION NUMBER: 60/227439
? PRIOR FILING DATE: 2000-08-22
? PRIOR APPLICATION NUMBER: MB1-0022
? PRIOR FILING DATE: 2001-11-16
? PRIOR APPLICATION NUMBER: MB1-0023
? PRIOR FILING DATE: 2001-04-17
? NUMBER OF SEQ ID NOS: 516
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 481
? LENGTH: 741
?
? TYPE: DNA
?
? ORGANISM: Arabidopsis thaliana
?
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(741)
?
? OTHER INFORMATION: G2422
?
? OS-09-934-455-481

```

Query Match	22.5%	Score 185.4	DB 11	Length 741
Best Local Similarity	74.3%	Pred. No. 1.4e-32		
Matches 234	Conservative 0	Mismatches 81	Indels 0	Gaps 0

Qy	28	GGAGTGGAAAAGGTCATGAGCATGATAGAAAAGATTTCTCTCAAGAAAATGTATTCAT	87
Db	19	GGGTTGAGAAAAGGTCATGGACTAGCTAGAGAGATATTTCTTGAGCCAAATGCAATATAT	78
Qy	88	AAGTATGGTGAAGGAAAATGGCATCTTGTTCCCAATAGAGCTGGTCTGAATAGATGTCCG	147
Db	79	AAGTATGGAGAGGCCAAATGGCATCGATTTCTTTAAGAACTGGTCTCAATCGGTGCCGA	138
Qy	148	AAAAGTTGTAGATTGAGGTGGCTGAATTTATCTAAGGCCACATATCAAGAGAGTGACTTT	207
Db	139	AAGAGTTGTAGACTTAAATGTTGTAATTTATTTGAAGCCAAAGATTTAAGAGGAAAACTC	198
Qy	208	GAAACAAGATGAAGTGGATCTCATATTTTGAAGCTTCAATAACTCTTAGSCAACAGATGTCA	267
Db	199	TGCTCCGATGAAGTGTGATCTTGTTCTTCGCTTCAATAACTTCTTAGGAAAAATGGGTGCC	258
Qy	268	CTTATTTGTGTAACTTCCCGGAAGACACTAAACATGTGAAAAACTATTTGAAACT	327
Db	259	TTGATCCGTGGTAAATGCCCTGGTCGACCTCTATATGATGTCAGAAATTACTGAAACT	318
Qy	328	AATCTTTAAGGAAG	342
Db	319	CATTGAGTAAAGG	333

```

: RESULT 8
: US-09-938-842A-2634
: Sequence 2634, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Krebs, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: FILE REFERENCE: SCRIPI300-3
: CURRENT APPLICATION NUMBER: US/09/938,842A
: CURRENT FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 2634
: LENGTH: 747
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana

```

US-09-938-842A-2634	
Query Match	21.7%;
Beet Local Similarity	72.1%;
Matches 233; Conservative	0; Mismatches 90; Indels 0; Gaps 0;

Qy	19	TCCTCATTGGGAGAGGAAAAAGGTTCTATGCACTGATGAGAGATTTTCTCTAAGAAA	78
Db	10	TCGTCAAAAGGGCTGCGAAAAAGGCTTGAGCTACTGAAAGAAAGATGCTCTTGAGACAG	69
Qy	79	TGTTATGATATAGTATGCTGAAAGAAAATGCGACTTGTCCCATTAAGACGTGCTGAAT	138
Db	70	TGCAATTAATATAGTATGAGGAAAGGCAAAATGGACCAAGTTCCTGTAAAGCTGGGCTAAAC	129
Qy	139	AGATGTGCGAAAAGTTGTAGATTGAGGTGGCTGTAATATCTTAAGGCCACATATCAAGAGA	198
Db	130	CGGTGACAGAAAAGTTGTAGATTAAAGATGGTTGAATATTTTGAAGCCCAAGTATCAAGAGA	189
Qy	199	GGTGACTTTGACAACAAGATGAAAGTGGATCTCAATTTGAGCGTTCACTAAGCTCTTAAAGCAC	258
Db	190	GGAAATCTGAGCTCTGATGAAAGTGGATCTTCTTTCGCTTCAATAGGCTTCTAAGGGAT	249
Qy	259	AGATGTCACATTATGCTGTAGACTTCCCGGAGAGACAGCTAAAGATGTGAAAAAATCAT	318
Db	250	AGGTGCTCTTAAATGCTGGAAGATATACGTGTGGACCGCAATGACGTCAAGAAATTAC	309
Qy	319	TGGAACAATAATCTTTAAGGAA	341
Db	310	TGGAACAATACTGAGTAAGAA	332

```

RESULT 9
US-09-938-842A-2634
; Sequence 2634, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIORITY FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2634
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2634

```

	Query Match	Similarity	21.7%	Score 179;	DB 12;	Length 747;
	Best Local	Similarity	72.1%	Pred. No. 4.2e-33;		
	Matches	233;	Conservative	0;	Mismatches 90;	Indels 0; Gaps 0;
Qy	19	TCCTCATTTGGAGATGAGAAAAAGTTTCATGCACTGATGAGAGATTTTCTTCTAAGAAAA	78			
Db	10	TCGTCCAAAGGGCGCGAAAAAGTGCTTTGAGCTACTGAAGAGATAGTCTCTTGAGACAG	69			
Qy	79	TGATTTGATAGTATGGTGAAGAAAAATGGCATCTTTGTTCCATTAAGCTGTGCTGAAT	138			
Db	70	TGCATTATTAATGATGAGAGAGCAAAATGGCACCAGTTCTCTTAAGCGCTGGGCTAAAC	129			
Qy	139	AGAGTGGGAAAAATGTTTGAATTGAGGTGGCTGAATATATCTAAAGCCACATATCAAGA	198			
Db	130	CGGTGCAGAAAAATGTTGATTAAGATGGTTGAACATATTGTAAGCCACATATCAAGA	189			



APPLICANT: Rafalski, Antoni  
APPLICANT: Shi, June  
APPLICANT: Weng, Zude  
TITLE OF INVENTION: Plant MYB-Related Transcription Factors  
FILE REFERENCE: B01280 US NA  
CURRENT APPLICATION NUMBER: US/09/443,704  
CURRENT FILING DATE: 1999-11-19  
EARLIER APPLICATION NUMBER: 60/109,294  
EARLIER FILING DATE: No. US2002066120A,ember 20, 1998  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 5  
LENGTH: 514  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (484)  
US-09-443-704-5

Query Match 19.1%; Score 157.4; DB 9; Length 514;  
Best Local Similarity 66.2%; Pred. No. 3.1e-26;  
Matches 227; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 9 TACATCTATGCTTTCATTTGGAGTGAAGAAAGTTCACTGATGAGAAAGATTTC 68  
DB 84 TCCTTGTGTTCCAAAGTGGGGTTGCACAAAGGTCATGAGCTCTTAAGAAATGCA 143  
QY 69 TCTAAGAAATGTAATGATTAAGTATGAGTGAAGAAATGSCATCTGTTCCATAAGAGC 128  
DB 144 GCTTACCAAGTATATCCAAAGCTCATGAGAGGCAATGAGAAATCACTACCAAAAAAGC 203  
QY 129 TGTCTGAATAGATGTCGAAAGTTGATGAGTGGCTGAATTAATTAAGGCCAC 188  
DB 204 AGGCTCTTGAAGTGGAAAGTTGATGAGTGAATGAATGAATCTTGAAGCCAGA 263  
QY 189 TATCAAGAGAGTGAATTTGAACAGATGAAGTGAATCTTATTTGAGCTTCAAGCT 248  
DB 264 TATTAAGAGAGGAACTATACCAAGAGAAAGTATTAATGATGATGATTCAC 323  
QY 249 CTTAAGCAAGATGCTCACTTATGCTGTGAACCTTCCGGAAGACAGTAAAGTGT 308  
DB 324 TTGGGAAACAGATGCTCCTCATAGCAGAGAGTTTACAGGAGAAACAGCAATGAAT 383  
QY 309 GAAAACTATTGGAACACTAATCTTCAAGAACTTAATACT 351  
DB 384 AAGAACTATTGGACACCCCATCTAAGCAAAAAAGCTGAATAAT 426

## RESULT 13

US-10-008-118A-5  
Sequence 5, Application US/10008118A  
Publication No. US20020187539A1  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Weng, Zude  
TITLE OF INVENTION: Plant MYB-Related Transcription Factors  
FILE REFERENCE: B01280 USDIY  
CURRENT APPLICATION NUMBER: US/10/008,118A  
CURRENT FILING DATE: 2001-12-05  
PRIOR APPLICATION NUMBER: 60/109,294  
PRIOR FILING DATE: 1998-11-20  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 5  
LENGTH: 514  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (484)  
OTHER INFORMATION: n = A, C, G or T  
US-10-008-118A-5

Query Match 19.1%; Score 157.4; DB 14; Length 514;  
Best Local Similarity 66.2%; Pred. No. 3.1e-26;  
Matches 227; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 9 TACATCTATGCTTTCATTTGGAGTGAAGAAAGTTCACTGATGAGAAAGATTTC 68  
DB 84 TCCTTGTGTTCCAAAGTGGGGTTGCACAAAGGTCATGAGCTCTTAAGAAATGCA 143  
QY 69 TCTAAGAAATGTAATGATTAAGTATGAGTGAAGAAATGSCATCTGTTCCATAAGAGC 128  
DB 144 GCTTACCAAGTATATCCAAAGCTCATGAGAGGCAATGAGAAATCACTACCAAAAAAGC 203  
QY 129 TGTCTGAATAGATGTCGAAAGTTGATGAGTGGCTGAATTAATTAAGGCCAC 188  
DB 204 AGGCTCTTGAAGTGGAAAGTTGATGAGTGAATGAATGAATCTTGAAGCCAGA 263  
QY 189 TATCAAGAGAGTGAATTTGAACAGATGAAGTGAATCTTATTTGAGCTTCAAGCT 248  
DB 264 TATTAAGAGAGGAACTATACCAAGAGAAAGTATTAATGATGATGATTCAC 323  
QY 249 CTTAAGCAAGATGCTCACTTATGCTGTGAACCTTCCGGAAGACAGTAAAGTGT 308  
DB 324 TTGGGAAACAGATGCTCCTCATAGCAGAGAGTTTACAGGAGAAACAGCAATGAAT 383  
QY 309 GAAAACTATTGGAACACTAATCTTCAAGAACTTAATACT 351  
DB 384 AAGAACTATTGGACACCCCATCTAAGCAAAAAAGCTGAATAAT 426

## RESULT 14

US-09-938-842A-2636  
Sequence 2636, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kreps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
FILE REFERENCE: SCRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 2636  
LENGTH: 774  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2636

Query Match 18.8%; Score 154.8; DB 10; Length 774;  
Best Local Similarity 66.5%; Pred. No. 1.5e-25;  
Matches 222; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 32 TGAGAAAGTTCATGAGCTGATGAAGAAATTTCTTAAGAAATGATTAAGT 91  
DB 35 TGAACAAAGAGCTTGAGCTTAAGAAAGATGAGCTTCTTGTGATTCATCCGTAAC 94  
QY 92 ATGTGAAGAAATGAGCATCTTGTCCCATTAAGAGCTGCTGAATAGATGCGAAAA 151  
DB 95 AGGTGAAGTGTGCTGGCATCTCTCCCGCGCGCTGATTAAGAAATGATGTAAG 154  
QY 152 GTTGTGATTTGAGTGGCTGAATTTATCTAAGCCATATCAAGAGAGTGAATTAAC 211  
DB 155 GTTGTGATTTGAGTGAATGAATTAATCTAAGCCAGATCTCAAAAGAGCAATTTACTG 214



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 06:37:27 ; Search time 1957 Seconds  
(without alignments)  
10245.883 Million cell updates/sec

Title: US-10-033-190-1  
Perfect score: 825  
Sequence: 1 atgaacagacatcatatgc.....tatgaatcactatgataa 825

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pla:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	253.2	30.7	490	10	BE462282 EST324546
2	252.2	30.6	418	9	AM626121 AM626121
3	222.6	27.0	335	10	BE462229 EST324493
4	196.8	23.9	749	13	BQ990780 BQ990780

5	185	22.4	367	10	BE635572 BE635572
6	184.4	22.4	486	9	AI995124 AI995124
7	172.2	20.9	427	13	BQ105368 BQ105368
8	172.2	20.9	488	13	BQ105048 BQ105048
9	172.2	20.9	508	13	BQ104423 BQ104423
10	165.8	20.1	551	12	BM092559 BM092559
11	165.6	20.0	494	14	CA410578 CA410578
12	165.4	19.1	587	14	H76020 H76020
13	157.4	18.8	397	12	BG881996 BG881996
14	155.2	18.8	781	13	BH830456 BH830456
15	154.8	18.8	636	13	BH827658 BH827658
16	154.6	18.7	447	10	BG134230 BG134230
17	154.6	18.7	574	14	CB088062 CB088062
18	153	18.5	529	9	AI727146 AI727146
19	153	18.5	620	9	AI730655 AI730655
20	151.8	18.4	447	9	AI486301 AI486301
21	151.8	18.4	562	9	AI771837 AI771837
22	151.6	18.4	505	14	CA902470 CA902470
23	151.4	18.4	522	12	BH557823 BH557823
24	151.4	18.4	617	9	AI730649 AI730649
25	151.4	18.4	642	12	BH573384 BH573384
26	151.4	18.4	556	9	AM686107 AM686107
27	151.4	18.4	797	13	BQ137572 BQ137572
28	150.6	18.3	355	9	AM278294 AM278294
29	150.6	18.3	646	14	CB917255 CB917255
30	150.2	18.2	671	13	BO511455 BO511455
31	150.2	18.2	757	12	BG889982 BG889982
32	149.6	18.1	572	9	AM559377 AM559377
33	149	18.1	475	9	AM186273 AM186273
34	148.2	18.0	671	10	BG447824 BG447824
35	147.6	17.9	737	13	BH891795 BH891795
36	147	17.8	468	14	CB922741 CB922741
37	146.6	17.8	852	9	AU236989 AU236989
38	145.4	17.6	957	10	BF269948 BF269948
39	144.4	17.5	548	12	BH563359 BH563359
40	143.6	17.4	505	12	BM528882 BM528882
41	143.4	17.4	645	13	BH816560 BH816560
42	142.8	17.3	549	13	BH812564 BH812564
43	142.6	17.3	721	10	BG123386 BG123386
44	142.4	17.3	491	14	CD480267 CD480267
45	142.2	17.2	633	14	CA902489 CA902489

#### ALIGNMENTS

RESULT 1  
LOCUS BE462282  
DEFINITION EST324546 tomato flower buds 0-3 mm, Cornell University  
KEYWORDS Lycopersicon esculentum CDNA clone cTOA12C2, mRNA sequence.  
ACCESSION BE462282  
VERSION BE462282.1 GI:9508051  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum

REFERENCE  
AUTHORS van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E., Jiang, F., Hansen, T., Craven, M.B., Bowman, C.L., Rinning, C.M., Niernan, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
Generation of ESTs from tomato flower tissue, 0-3 mm buds

TITLE  
JOURNAL  
COMMENT Unpublished  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES  
source  
1. 490  
Location/Qualifiers

```

/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CT0A12C2"
/tissue_type="flower"
/dev_stage="0-3mm buds"
/clone_1ib="tomato flower buds 0-3 mm, Cornell University"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Tanksley; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

BASE COUNT      156 a      93 c      105 g      136 t
ORIGIN

Query Match      30.7%; Score 253.2; DB 10; Length 490;
Best Local Similarity 82.2%; Pred. No. 4.2e-39;
Matches 291; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1 ATGAACAGTACATCATGTCTTCATGTGGAGTGAGAAAGTTTCATGACTGATGAAGAA 60
DB 3 ATGAATATCTCTATGTGTGCATCTTGAGATTGGAAGTTTCATGACTGATGAAGAA 62
QY 61 GATTTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 63 GATTTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 122
QY 121 ATAAAGCTGTCTGAATAGATGTCGAAAGTTTGATGATGATGATGATGATGATGAT 180
DB 123 GCTAGAGCTGTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 182
QY 181 AGGCCATATATCAAGAGAGTGAATTGAAACAAGATGATGATGATGATGATGATGATGAT 240
DB 183 AGGCCATATATCAAGAGAGTGAATTGCTCCAGATGATGATGATGATGATGATGATGAT 242
QY 241 CATAGCTCTTAAGCAACAGATGTCATTTATGCTGTGATGATGATGATGATGATGATGAT 300
DB 243 CACAACTCTTAAGCAACAGATGTCATTTATGCTGTGATGATGATGATGATGATGATGAT 302
QY 301 AACGATGTGAAAACTTTGGACACTAATCTTTTAAAGAAAGTTAAATACTACT 354
DB 303 AACGATGTGAAAACTTTGGACACTAATCTTTTAAAGAAAGTTAAATACTACT 356

RESULT 2
AM626121      418 bp      mRNA      linear      EST 18-MAY-2001
LOCUS      EST320028 tomato radicle, 5 d post-imbibition, Cornell University
DEFINITION      Lycopersicon esculentum cDNA clone cLEZ18K18 5', mRNA sequence.
ACCESSION      AM626121
VERSION      AM626121.1 GI:7339148
KEYWORDS
SOURCE      EST.
ORGANISM      Lycopersicon esculentum (tomato)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 418)
REFERENCE      van der Hoeven, R.S., Garvin, D., Matern, A.L., Holt, I.E., Liang, F.,
AUTHORS      Hansen, T., Craven, M.B., Bowman, C.L., Romning, C.M., Nierman, W.,
Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
Generation of ESTs from tomato radicle tissue (ecoliated)
JOURNAL      Unpublished
COMMENT      Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1..418
/organism="Lycopersicon esculentum"

```

```

/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEZ18K18"
/tissue_type="radicle"
/dev_stage="seedlings 5 days post-imbibition"
/clone_1ib="tomato radicle, 5 d post-imbibition, Cornell University"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Tanksley; Tissue supplied by Dave Garvin (USDA-ARS, Ithaca, NY 14850)."

BASE COUNT      128 a      79 c      91 g      119 t      1 others
ORIGIN

Query Match      30.6%; Score 252.2; DB 9; Length 418;
Best Local Similarity 81.9%; Pred. No. 6.5e-39;
Matches 290; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 1 ATGAACAGTACATCATGTCTTCATGTGGAGTGAGAAAGTTTCATGACTGATGAAGAA 60
DB 3 ATGAATATCTCTATGTGTGCATCTTGAGATTGGAAGTTTCATGACTGATGAAGAA 62
QY 61 GATTTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 63 GATTTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 122
QY 121 ATAAAGCTGTCTGAATAGATGTCGAAAGTTTGATGATGATGATGATGATGATGAT 180
DB 123 GCTAGAGCTGTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 182
QY 181 AGGCCATATATCAAGAGAGTGAATTGAAACAAGATGATGATGATGATGATGATGATGAT 240
DB 183 AGGCCATATATCAAGAGAGTGAATTGCTCCAGATGATGATGATGATGATGATGATGAT 242
QY 241 CATAGCTCTTAAGCAACAGATGTCATTTATGCTGTGATGATGATGATGATGATGATGAT 300
DB 243 CACAACTCTTAAGCAACAGATGTCATTTATGCTGTGATGATGATGATGATGATGATGAT 302
QY 301 AACGATGTGAAAACTTTGGACACTAATCTTTTAAAGAAAGTTAAATACTACT 354
DB 303 AACGATGTGAAAACTTTGGACACTAATCTTTTAAAGAAAGTTAAATACTACT 356

RESULT 3
BE462229      335 bp      mRNA      linear      EST 18-MAY-2001
LOCUS      EST324493 tomato flower buds 0-3 mm, Cornell University
DEFINITION      Lycopersicon esculentum cDNA clone cTOA12C3, mRNA sequence.
ACCESSION      BE462229
VERSION      BE462229.1 GI:9507998
KEYWORDS
SOURCE      EST.
ORGANISM      Lycopersicon esculentum (tomato)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 335)
REFERENCE      van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E., Liang,
AUTHORS      F., Hansen, T., Craven, M.B., Bowman, C.L., Romning, C.M., Nierman, W.,
Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
Generation of ESTs from tomato flower tissue, 0-3 mm buds
JOURNAL      Unpublished
COMMENT      Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1..335
/organism="Lycopersicon esculentum"

```



```

/db_xref="taxon:4081"
/clone="CT0A12C3"
/cisue_type="flower"
/dev stage="0-3mm buds"
/clone lib="tomato flower buds 0-3 mm, Cornell University"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Tankley; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."
BASE COUNT      109 a      54 c      77 g      95 t
ORIGIN
Query Match      27.0%; Score 222.6; DB 10; Length 335;
Best Local Similarity 80.3%; Pred. No. 3.5e-33;
Matches 261; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 1 ATGACAGTACATCTATGCTTCATTTGGAGTGAAGAAAGTTGATGATGAGAA 60
DB 11 ATGAAATCTCTATGCTGATGCTTGGAGTTAGAGAAAGTTGATGATGAGAA 70
QY 61 GATTTCTTCTAGAAATGATGATGATGATGATGATGATGATGATGATGATG 120
DB 71 TATTTCTTCTAGAAATGATGATGATGATGATGATGATGATGATGATGATG 130
QY 121 ATGAGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 180
DB 131 GCTACAGTGTGTTAAATCTATGCTGAAAGTTGATGATGATGATGATGAT 190
QY 181 AGCCCATATGACAGAGAGTGAAGTTGAAAGATGATGATGATGATGATGAT 240
DB 191 AGCCCATATGACAGAGAGTGAAGTTGAAAGATGATGATGATGATGATGAT 250
QY 241 CATTAGCTCTAGGCAACAGATGATGATGATGATGATGATGATGATGATGAT 300
DB 251 GACAACTCTTATGACATAGCGGTCACTTATGCTGATGATGATGATGATG 310
QY 301 AACGATGTGAAAGATGATGATGATGATGATGATGATGATGATGATGATG 325
DB 311 AACGATGTGAAAGATGATGATGATGATGATGATGATGATGATGATGATG 335

RESULT 4
BO990780      749 bp      mRNA      linear      EST 21-AUG-2002
LOCUS      OGF21B10.yg.ab1 OG_EFGHJ lettuce serritola Lactuca sativa cDNA clone
DEFINITION      OGF21B10, mRNA sequence.
ACCESSION      BO990780
VERSION      BO990780
KEYWORDS      EST.
SOURCE      BO990780.1 GI:22410315
ORGANISM      Lactuca sativa
                Lactuca sativa
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
                Cichoriaceae; Lactuca.
                1 (bases 1 to 749)
                Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
                Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison
                ,P., Kojkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
                Church,S., Jackson,L. and Bradford,K.
                lettuce and Sunflower ESTs from the Composite Genome Project
                http://compgenomics.ucdavis.edu/
                Unpublished
                Contact: Alexander Kozik [R.W.Michelmore]
                Department of Vegetable Crops, R.W.Michelmore Lab
                University of California at Davis (UCD)
                Aemundson Hall, UCD, Davis, CA 95616, USA
                Tel: 1-(530)-742-1742
                Fax: 1-(530)-752-3659
                Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
                singleton, see http://cgpdb.ucdavis.edu/ for details.
                Plate: OGF21 row: B column: 10.

```

```

FEATURES
    source
        Location/Qualifiers
            1..749
                /organism="Lactuca sativa"
                /mol_type="mRNA"
                /cultivar="L.serritola"
                /db_xref="taxon:4236"
                /clone="OGF21B10"
                /lab_host="E.coli"
                /clone lib="OG_EFGHJ lettuce serritola"
                /note="Vector: pBRDNASf1AB; The library was constructed
                from 10 different sources of RNA from a single genotype.
                Separate cDNAs were generated using primers that
                incorporated unique 5' and 3' tags to distinguish each
                source of RNA. cDNAs were then pooled, size-fractionated,
                directionally cloned into a custom medium-copy vector and
                transformations made with four size classes to minimize
                size bias. Details of each source of RNA and library
                construction can be obtained at http://cgpdb.ucdavis.edu/
                TAG LIB=OG_EFGHJ lettuce serritola
                TAG TISSUE=leaves dark grow
                TAG SRO=GTAGTGGG"
BASE COUNT      266 a      129 c      157 g      196 t      1 others
ORIGIN
Query Match      23.9%; Score 196.8; DB 13; Length 749;
Best Local Similarity 75.0%; Pred. No. 3.6e-28;
Matches 246; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 9 TACATCTATGCTCTTCATGAGAGTGAAGAAAGTTGATGATGATGATGATGAT 68
DB 142 TACTACTACTCATGCTTGTAGATTAAGAAAGTGCATGATGATGATGATGAT 201
QY 69 TCTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 128
DB 202 TCTTAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 261
QY 129 TGGTCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 188
DB 262 AGGGTTAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 321
QY 189 TATCAAGAGAGTGAAGTTGAAAGATGATGATGATGATGATGATGATGATGAT 248
DB 322 TATTAAGAGAGAAATTTGTGATGATGATGATGATGATGATGATGATGATGAT 381
QY 249 CTTAGGCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 308
DB 382 TTTAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 441
QY 309 GAAAACTATTGGAACATGATGATGATGATGATGATGATGATGATGATGATG 336
DB 442 GAAAGACTACTGGAACAAATTTTCAA 469

RESULT 5
BF635572      367 bp      mRNA      linear      EST 19-DEC-2000
LOCUS      NF104H01D1.F1014 Drought Medicago truncatula cDNA clone NF104H01D
DEFINITION      5', mRNA sequence.
ACCESSION      BF635572
VERSION      BF635572
KEYWORDS      EST.
SOURCE      BF635572.1 GI:11899730
ORGANISM      Medicago truncatula (barrel medic)
                Medicago truncatula
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
                ; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
                Medicago.
                1 (bases 1 to 367)
                Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
                Flores,H.R., Inman,U.T., Weller,U.W. and May,G.D.
                Expressed Sequence Tags from the Samuel Roberts Noble Foundation
                Medicago truncatula drought library
                Unpublished

```

COMMENT Contact: May GD  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org  
Insert Length: 367 Std Error: 0.00  
Plate: 104 row: H column: 01  
Seq primer: TCACACAGAAACGCTATGAC.  
Location/Qualifiers

FEATURES  
source  
1. 367  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/db\_xref="taxon:3880"  
/clone="NF104H01DT"  
/issue\_type="Plantlets"  
/dev\_stage="Pooled timepoints"  
/clone\_1lb="Drought"  
/note="Vector: Lambda Zap; Contains a mixture of entire plantlets harvested in a series of days-post-watering timepoints."  
BASE COUNT 142 a 50 c 85 g 86 t 4 others  
ORIGIN

Query Match 22.4%; Score 185; DB 10; Length 367;  
Best Local Similarity 73.3%; Pred. No. 6.7e-26;  
Matches 236; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 36 AAAAGGTCATGACATGATGAGAGATTTCTTCTAAGAAATGATATGATATG 95  
DB 1 AAAAGGTCATGACATGATGAGAGATTTCTTCTAAGAAATGATATGATATG 60  
QY 96 TGAAGGAAATGCGATCTTGTCCATTAAGAGCTGCTGAATGATGCGAAAAATTG 155  
DB 61 TGAAGGAAATGCGATCTTGTCCATTAAGAGCTGCTGAATGATGCGAAAAATTG 120  
QY 156 TAGATGAGGCTGCTGAATTTCTTAAGGCCATATCAAGAGAGTGAACAGAGA 215  
DB 121 TAGATGAGGCTGCTGAATTTCTTAAGGCCATATCAAGAGAGTGAACAGAGA 180  
QY 216 TGAAGTGCATCTCATTTTGAAGCTTCTTAAGGCCAAGATGCTCATTTATGC 275  
DB 181 TGAATTTGATATGATTTCTAAGGTACACAACTTCTAGGGAACAAATGCTTTGATTCG 240  
QY 276 TGTGAGACTTCCCGAAGACAGCTAAGATGTAAAAACATTTGAACTATCTTCT 335  
DB 241 TGCAGAGCTTCCCGGTAGGACAGCTAATGATGAAANAATTTGGACACAAATTTGCG 300  
QY 336 AAGGAGTTAAATCTACTATA 357  
DB 301 CAAAAAGTATTCAGAGAAA 322

RESULT 6  
AI995124 486 bp mRNA linear EST 08-SEP-1999  
LOCUS 701502113 A. thaliana, Ohio State clone set Arabidopsis thaliana  
DEFINITION cDNA clone 701502113, mRNA sequence.  
ACCESSION AI995124  
VERSION AI995124.1 GI:5842029  
KEYWORDS EST.  
ORGANISM Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroids I; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 486)  
Chen, J., Montyana, M., Chan, E., Mooney, M., Caroon, B., Gilliland, D., Wang, X., Hillman, J., Giegler, K., Kim, C., Doyle, M., Brzoka, P., Gorgone, G., Burns, D., Griffin, J., McAnoutou, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kascary, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nodrig, A., Murry, L.,

TITLE Turner, C., Krikorian, S., Elder, L. and Hanson, D.  
JOURNAL Arabidopsis thaliana Gene Expression Microarray  
COMMENT Unpublished  
Contact: David Smoller, Ph.D.  
Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.  
4633 World Parkway Circle, St. Louis, MO 63134, USA  
Tel: 877-577-2733  
Fax: 314-427-3324  
Email: service@genomesystems.com.  
Location/Qualifiers

FEATURES  
source  
1. 486  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="701502113"  
/clone\_1lb="A. thaliana, Ohio State clone set"  
/note="cDNA library was made from selected clones from the Arabidopsis thaliana Ohio State clone set."  
BASE COUNT 148 a 83 c 111 g 144 t  
ORIGIN

Query Match 22.4%; Score 184.4; DB 9; Length 486;  
Best Local Similarity 74.2%; Pred. No. 9e-26;  
Matches 233; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 19 TCTTCATTTGGAGATGAGAAAGGTCATGACATGATGAGAGATTTCTTCTAAGAAA 78  
DB 105 TCGTCCAAAGGCTTGGAGAAAGTGCATGACCTGAGAGATATGCTCTTGAAGCTA 164  
QY 79 TGTATTGATTAAGATGATGAGAAATGCGATCTTGTTCCTAATGAGCTGTGTAAT 138  
DB 165 TGTATTGATTAAGATGAGAAAGGCAATGCGATCAAGTCTTTGAGAGCTGGCTAAAT 224  
QY 139 AGATTCGAGAAAGTTGTATGATGAGCTGCTGAATTTATCTAAGGCCATATCAAGAGA 198  
DB 225 CGATGAGAAAGGTTGTATGATGAGCTGCTGAATTTATCTAAGGCCATATCAAGAGA 284  
QY 199 GGTGACTTTGAACAAAGATGAGATGATCTATTTGAGGCTTCATTAAGCTTGAAGCAAC 258  
DB 285 GGAAGACTTGAACAAAGATGAGATGATCTATTTGAGGCTTCATTAAGCTTGAAGCAAC 344  
QY 259 AGATTCGATCTTATTTCTGTAGACTTCCCGAAGACAGCTAAGCTGAAAACTAT 318  
DB 345 AGGTGGCTCTTATTTCTGTAGACTTCCCGAAGACAGCTAAGCTGAAAACTAT 404  
QY 319 TGGACACTATATCT 332  
DB 405 TGGACACCCATCT 418

RESULT 7  
B0105368 427 bp mRNA linear EST 16-APR-2002  
LOCUS fc0707.e Rose Petals (Fragrant Cloud) Lambda Zap Express Library  
DEFINITION Rosa hybrid cultivar cDNA clone fc0707.e 5', mRNA sequence.  
ACCESSION B0105368  
VERSION B0105368.1 GI:20155030  
KEYWORDS EST.  
ORGANISM Rosa hybrid cultivar  
Rosa hybrid cultivar  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroids I; Rosales; Rosaceae; Rosoideae; Rosa.  
1 (bases 1 to 427)  
Guterman, I., Shalit, M., Menda, N., Piestun, D., Dafny-Yelin, M., Shalev, G., Bar, E., Davidov, O., Ovadis, M., Emanuel, M., Wang, J., Adam, Z., Pichersky, E., Lewinsohn, E., Zamir, D., Vainstein, A. and Weiss, D.  
Rose Scent: Genomics Approach to Discovering Novel Floral Fragrance-Related Genes  
Plant Cell 14 (10), 2325-2338 (2002)  
Contact: Naama Menda

**Petal Genomics**  
Faculty of Agricultural, Food and Environmental Quality Sciences,  
The Hebrew University of Jerusalem  
P.O. Box 12, Rehovot, 76100, Israel  
Tel: 972 8 9489 389  
Fax: 972 8 9468 263  
Email: shaham@agri.huji.ac.il  
Seq primer: 73 forward, 64 reverse

BASE COUNT	142 a	82 c	111 g	92 t	-
ORIGIN					

Query Match	20.9%;	Score 172.2;	DB 13;	Length 427;
Best Local Similarity	67.3%;	Pred. No. 2.1e-23;		
Matches 243;	Conservative 0;	Mismatches 118;	Indels 0;	Gaps 0;

QY	42	TTCAATGACTGATGAAAGAGTTTCTTCTTAAGAAATGTATATGATAGTGTAAAGG	101
Db	14	TTCAATGACTAAGGAAGAGATCAATTGTCTGAAGAAATGCTATAGAAAGATGTGAAGG	73
QY	102	AAATATGACATCTTGTTCCCATTAAGCTGTCTCTGAATAGATGTGGAAAAATTGTAGATT	161
Db	74	CAAAATGGACATCGAATTTCTCATCTTAGCCGGTCTTAACAGATGCGTAAAGAGCTCCGGCT	133
QY	162	GAGTGTGCTGAATTTATCTTAAGCCACATATCAAGAGAGGTGACTTTTGAAACAAATGAAGT	221
Db	134	TCGGTGGCTTAAACTACCTTCGCGCCGAACATCAAGAGAGGAGCGTTTGAACAAAGAGAGT	193
QY	222	GGATCTCATTTTGGAGCTTCATTAAGCTCTTAGGCAACAGATGCTCATTAATGCTGGTAG	281
Db	194	TGAGCTCATCATCAAGACTCCAAAGCTCTTTGGGAAACAGGTGTCTATTAATTGCAGCTAG	253
QY	282	ACTTCGCCGAAGAGACGCTAACGATGTGAAAACTATTGGAAACACTAATCTTCTTAAGAA	341
Db	254	GCTTCGCGGGAGAGACTGTAAATGATGTGAAGAACTATTGGAAATGTGCATGTGACAAAA	313
QY	342	GTTAAATACTACTAAATATTGTTCTTCGCGAAAAAGATTAACATTAAGTGTGGAGAAATTAG	401
Db	314	GCTAATATGCTCAAGAAACTCATGTGTATCATCAGAAAGATTAATCATGACAGGGAAATATAA	373
QY	402	T 402	
Db	374	T 374	

RESULT 8	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
BO105048	BO105048	488 bp mRNA linear EST 16-APR-2002	g91664.e	Rose Petals (Golden Gate) Lambda Zap Express Library Rosa hybrid cultivar cDNA clone g91664.e 5', mRNA sequence.	BO105048.1	GI:20154710	EST.	Rosa hybrid cultivar
								Rosa hybrid cultivar
								Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
								; eusoids I; Rosales; Rosaceae; Rosoideae; Rosa.
								1 (bases 1 to 488)
								Guerman, I., Shalit, M., Menda, N., Pietsch, D., Dafny-Yelin, M., Shalev, G., Bar, E., Davydov, O., Ovadis, M., Emanuel, M., Wang, J., Adam, Z., Pitchersky, E., Lewinsohn, E., Zamir, D., Vainstein, A. and Weiss

TITLE	,D. Rose Scent: Genomics Approach to Discovering Novel Floral Fragrance-Related Genes
JOURNAL	Plant Cell 14 (10), 2325-2338 (2002)
COMMENT	Contact: Naama Menda Plant Genomics

BASE COUNT	161 a	92 c	128 g	107 t
ORIGIN				

Query Match	20.9%	Score 172.2;	DB 13;	Length 488;
Best Local Similarity	67.3%	Pred. No. 2.1e-23;		
Matches 243; Conservative	0;	Mismatches 118;	Indels 0;	Gaps 0;

Qy	4	TTTCATGACCTGATGAAGAAGATTTCTTCTTAAGAAAAATGTAATGATTAAGATGCTGAAG	101
Db	22	TCCCTGGACTTAAGAAAGAGACAGTCTGCTAGGAATGCATAGAGAATGTGAAGAAG	81
Qy	102	AAATGGACCTTGTGTTCCCATAGAGCTGTCTGAATAGATGTCCGAAAAGTTGTAGATT	161
Db	82	CAATGGACCTGAATTTCTCACTTAGCCGGCTTAAGACGGGCCGTAAAGAGCTGCAGGCT	141
Qy	162	GAGGTGGCTGAATTATCTTAAGGCCACATATCAAGAGAGTGACTTTTAAACAATGAAGT	221
Db	142	TGCGTGGCTAAACCTACCTTCGGCCGAACATMAAGAGGGAGCTTTGAACAAGAGGAAGT	201
Qy	222	GGATCTCATTTTGAAGGCTTCAATAGCTCTTAAGGCACAAGATGGTCATTAATGCTGGTAG	281
Db	202	TGAGCTCATCATCCAGCTCCAAAGGCTTTGGAAAAAGGTGGTCATTAGTTGACAGTAG	261
Qy	282	ACTTCCCGAGAGCAGCTAAAGATGTGAAAAACTATTGGAACACTAATCTTCTTAAGAA	341
Db	262	GCCTCCGGGAGAGACTGGTAATGATGTGAAGACTAATGGAAATGTCACTTGAGCAAAAA	321
Qy	342	GTTTAATTACTAATAAATTTGTTCTCCGGAAAAAGTTAACAATTAAGTGGAGAAATTAG	401
Db	322	GCTAAATGCTCAAGAAACTATGGATCATCAAGAAAGTTAATCATGACAGGGAATATATA	381
Qy	402	T 402	
Db	382	T 382	

RESULT 9	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
B0104423	BQ104423	508 bp mRNA linear EST 16-APR-2002	g90628.e	Rose Petals (Golden Gate) Lambda Zap Express Library	Rosa	hybrid cultivar cDNA clone g90628.e 5', mRNA sequence.	<p>BQ104423</p> <p>BQ104423</p> <p>BQ104423.1 GI:20154085</p> <p>EST.</p> <p>Rosa hybrid cultivar</p> <p>Rosa hybrid cultivar</p> <p>Eucaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Magnoliopsida; Magnoliophyta; eudicotyledons; core eudicots; rosids</p>

REFERENCE 1 (bases 1 to 508)  
 AUTHORS Guterman, I., Shalit, M., Menda, N., Pleskun, D., Dafny-Yelin, M., Shalev, G., Bar, E., Dayadov, O., Ovadis, M., Emanuel, M., Wang, J., Adam, Z., Pichersky, E., Lewinsohn, E., Zamir, D., Vainstein, A. and Weiss, D.  
 TITLE Rose Scent: Genomics Approach to Discovering Novel Floral Fragrance-Related Genes  
 JOURNAL Plant Cell 14 (10), 2325-2338 (2002)  
 COMMENT Contact: Naama Menda  
 Petal Genomics  
 Faculty of Agricultural, Food and Environmental Quality Sciences,  
 The Hebrew University of Jerusalem  
 P.O. Box 12, Rehovot, 76100, Israel  
 Tel: 972 8 9489 389  
 Fax: 972 8 9468 263  
 Email: snaham@agri.huji.ac.il  
 Seq primer: T3 forward.  
 FEATURES  
 source  
 1. 508  
 Location/Qualifiers  
 /organism="Rosa hybrid cultivar"  
 /mol\_type="mRNA"  
 /strain="Golden Gate"  
 /db\_xref="taxon:128735"  
 /clone="g90628.e"  
 /tissue\_type="Petals"  
 /dev\_stage="Young open flower at stage four"  
 /clone\_lib="Rose Petals (Golden Gate) Lambda Zap Express Library"  
 /note="Vector: pBKCW; Site 1: EcoRI; Site 2: XhoI"  
 BASE COUNT 170 a 92 c 137 g 109 t  
 ORIGIN  
 Query Match 20.9%; Score 172.2; DB 13; Length 508;  
 Best Local Similarity 67.3%; Pred. No. 2.1e-23;  
 Matches 243; Conservative 0; Mismatches 118; Indels 0; Gaps 0;  
 42 TTCATGACGTGATGAGAGATTTCTTCTAAGAAATGTATGATAGTATGTTGAAGG 101  
 Db TCCTAGGACTTAAAGAGATGATCAGTCTGCTAGGAAATGCTAAGAGATGAGAGAGG 90  
 QY 102 AAAATGGCATCTTGTTCCTCATAGAGCTGTCTGATATGATGTGCGAAAAGTTGTAGATT 161  
 Db 91 CAATGGATGGAATTCCTCAGTACGCGGTCTAAACAGGTGCGGTAAAGCTCAGGCT 150  
 QY 162 GAGGTGCTGAATTAATCTAAGGCCACATATCAAGAGAGTACTTTGAACAAGTGAAGT 221  
 Db 151 TCGGTGGCTAAACTACCTTCGCGCAACATCAAGAGAGGAGCTTTGAACAAGAGAACT 210  
 QY 222 GGATCTCAATTTGAGGGCTTCAAGCTCTTAGGCAACAGATGTCACCTATTGCTGGTAG 281  
 Db 211 TGAGCTATCATCCAGCTCCAAAGGCTTTGGAAACAGGTGTCATTAAGTTCAGACTAG 270  
 QY 282 ACTTCCCGAAGAGACAGCTAAGATGTGAAAACTATTGGAACATACTTCTTAAGGAA 341  
 Db 271 GCTTCGGGAGAGAGCTGTATATGATGAAGAACTATTGATGCTCAGACCAAAA 330  
 QY 342 GTTAATATCTAATAATTTGTTCTCGGAAAAGATTACATTAAGTGTGAGAAATTAG 401  
 Db 331 GCTAAATGCTCAAGAACTCATGTGATCATCAGAAATTAATCATGAGGGAATTAATA 390  
 QY 402 T 402  
 Db 391 T 391  
 RESULT 10  
 BM092559 551 bp mRNA linear EST 29-NOV-2001  
 LOCUS BM092559  
 DEFINITION ba15d11.y3 Gm-cl086 glycine max cdna clone GENOME SYSTEMS CLONE  
 ID: Gm-cl086-1342 5' similar to TR:023891 O23891 OSMYB3.1; mRNA  
 Sequence.  
 ACCESSION BM092559

VERSION BM092559.1 GI:17021525  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 551)  
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvett, V., Khanna, A., Bolla, B., Matra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Treising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished  
 COMMENT Contact: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: cu@resgen.com  
 High quality sequence stop: 421.  
 FEATURES  
 source  
 1. 551  
 Location/Qualifiers  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl086-1342"  
 /tissue\_type="Young seeds (Williams 82)"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-cl086"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from very young seeds (less than 20mg). The library was prepared using the Stratagene pBluescript II SK (+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(AT) sequence with an xho I restriction site. Eco RI adaptors were ligated to the blunt-ended cDNA fragments followed by xho I digestion. The cDNA insert is protected from xho I digestion via methylation during first strand cDNA synthesis. The cDNA fragments were directionally cloned into the Eco RI-Xho I restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cell. The library was constructed by Ann Khanna (Ulla Vodkin lab, University of Illinois)."  
 BASE COUNT 187 a 127 c 119 g 118 t  
 ORIGIN  
 Query Match 20.1%; Score 165.8; DB 12; Length 551;  
 Best Local Similarity 67.5%; Pred. No. 3.6e-22;  
 Matches 233; Conservative 0; Mismatches 112; Indels 0; Gaps 0;  
 3 GAACAGTACATCTATGCTTCATTTGGAGAGTGAAGAAAGTTCTATGACAGTGAAGA 62  
 Db 4 GAGAGGCCCTGTGTTCAAGAGAGGTTTGAATAGAGGCTTGGACAGCTCATGAAGA 63  
 QY 63 TTTTCTTAAGAAATGTATGATTAAGTGTGAAGAAATGCGCATTTGTTCCCAT 122  
 Db 64 CAAATCTCTCAGAGATACATTAGATCCATGTGGAAGAGATGAGAAACCTTCCAA 123  
 QY 123 AAGAGCTGTCTGAATGATGATGCGAAAGTGTAGATTGAGGAGGCTGAATATCTAAG 182  
 Db 124 AAGAGCAGGTTTGAAGAGATGAGAAAGTTGACAGCTTAAGTGTGTAATATCTCAG 183  
 QY 183 GCCACATATCAAGAGAGTGAATTTGAACAAGATGAAGTGAATCTCATTTTGAAGCTTCA 242

Db	184	ACGAGATATTAAAGAGGCATATATATCCCGAGATGAAGAAGCTCATTCAGGCTACA	243
Oy	243	TAACTTTAGGCCAACAGATGGTCACTTATTTGCTGTAGACTTCCCAGAGACAGCTAA	302
Db	244	CMACTCTCTGGAAAACAGATGGTCTTTTAATTAGCTGGGGGGCTTCAGAGACGACAGACAA	303
Oy	303	CGATGTGAAAAAAGCTATTGGAAACATACTTCTTAAGAGAGTTAA	347
Db	304	TGAATAATAAGATTAATTGGAAACCACCAATTATGAGAAAAAGGTGAA	348
 RESULT 11			
LOCUS	CA410578	494 bp	mRNA linear EST 01-JAN-2003
DEFINITION	843_F -P proteoid roots 12 and 14 DAE Lupinus albus cDNA clone 843		
ACCESSION	CA410578		
VERSION	CA410578.1	GI:27459582	
KEYWORDS	EST.		
SOURCE	Lupinus albus (white lupine)		
ORGANISM	Lupinus albus		
REFERENCE	1 (bases 1 to 494)		
AUTHORS	Unde-Stone,C., Zimm,K.E., Ramirez-Yanez,M., Li,A., Vance,C.P. and Allan,D.L.		
TITLE	Nylon filter arrays reveal differential gene expression in proteoid roots of white lupin in response to P deficiency		
JOURNAL COMMENT	Unpublished (2003) Contact: Unde-Stone, Claudia Department of Agronomy and Plant Genetics University of Minnesota 439 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA Tel: 612 624 6765 Fax: 612 625 2208 Email: cunde@soils.umn.edu PCR Primers FORWARD: T3 BACKWARD: T7 Seq primer: AATTAACTCCTCACTRAAGGG. Location/Qualifiers 1..494 /organism="Lupinus albus" /mol_type="mRNA" /cultiivar="vultura" /db_xref="taxon:3870" /clone="843" /dev_stage="12 and 14 DAE" /note=Organ: -P proteoid root (cluster roots); Vector: pBUESCRIP SK(+/-), Site_1: EcoRI; Site_2: XhoI; Proteoid roots of white lupin grown under P deficient conditions were harvested 12 and 14 DAE (days after emergence), and poly(A)+ RNA was isolated. The poly(A)+ RNA obtained from plants 12 and 14 DAE was combined in a 1:1 ratio and 7 ug total RNA was used for reconstruction of a proteoid root cDNA library in the phage ZapR vector according to the manufacturer's instructions (Stratagene, La Jolla, CA, USA). For conversion of the phage library into the plasmid form, mass excision was performed according to the procedure described by Stratagene."		
FEATURES	source		
BASE COUNT	197 a	71 c	111 g 96 t 19 others
ORIGIN			
Query Match	20.1%	Score 165.6;	DB 14; Length 494;
Best Local Similarity	63.7%;	Pred. No. 3.9e-22;	
Matches 249;	Conservative 0;	Mismatches 142;	Indels 0; Gaps 0;
19	TCTTCATTGGAGTGAGAAAGGTTCACTGAGCTGATGAAGAAGATTTCTTCTTAAGAAA	78	

Db		44	TCTAAAGTGGGTTGCACAAAGGTCANVGACTCTCTAAAGAAGATGATTACTTAACCAA	103
OY		79	TGTATTGAATAAGTATGGTAGAAGAAAATGGCATTTGTTCCCATPAAGAGCTGGTCTGAAT	138
Db		104	TATATTCAAGTCTCAYGAGAAGAGCGCAATGTGMAATCACCTAACCAAAAAAGGCTGGCTTCTT	163
OY		139	AGATGTGGSAAGAAATTGTAGATTGAGGTGGCTGAATTAATCTPAAAGGCCACATATCAAGAGA	198
Db		164	AGATGTGAAAAAATTGTAGATTGAGGTGGTAGTAATCTATCAAGACCAAGATATPAAAGAGA	223
OY		199	GGTGACTTTTGAAACAAGATGAAAGTGAATCTCATTTTGGAGCTTCAATAGCTCTTAGGCAAC	258
Db		224	GGTAAACATTAATCCAGAGAAGAAAGATCTCATCTCAAGATGCAATTTCTCTTTGGAAAT	283
OY		259	AGATGTGCTACTTATTTGCTGTGTAGACTTCCCGGAAGACAGCTPAAAGATGTGAAAACTAT	318
Db		284	AGGTGATCCCTTATPAGCCGGAAGAAATCCCGSAGACGATATATGAATPAAAGACTAT	343
OY		319	TGGAACACTAATCTCTTAAGGAAGTTAAATPACTTAAATTTGTTCTCCGGAAAGATTT	378
Db		344	TGGAACACCCCACTTTATACAAAAAGATGATAAGAAACAGATCTPAGAAAGAAATCT	403
OY		379	AACATPAAAGTGTGAGAAATTAATGACTPAGA	409
Db		404	CCCACTGTCAACAAAAAGANNCCAAGAAAG	434
RESULT 12				
H76020		587 bp	mRNA	linear EST 30-DEC-1997
LOCUS	17725	Lambda-PRL2	Arabidopsis thaliana	cDNA clone 193M157, mRNA sequence.
DEFINITION	H76020			
ACCESSION	H76020			
VERSION	H76020.1			GI:1053271
KEYWORDS	EST.			
SOURCE	Arabidopsis thaliana (thale cress)			
ORGANISM	Arabidopsis thaliana			
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids II; Brassicales; Brassicaceae; Arabidopsids.			
AUTHORS	Newman,T., deBrenin,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohltge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E., and Somerville,C.			
TITLE	Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)			
JOURNAL MEDLINE PUBMED COMMENT	95148729 7846151 Contact: Thomas Newman MSU-DOE Plant Research Laboratory Michigan State University MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI Tel.: 517-353-0854 Fax: 517-353-9168 Email: 22313tc@ibm.cl.msu.edu Seq primer: T7 dye primer. Location/Qualifiers 1..587 /organism="Arabidopsis thaliana" /mol_type="mRNA" /strain="varolumbia" /db_xref="taxon:3702" /clone="193M157" /clone_1b="Lambda-PRL2" /note="Vector: Lambda Zip-Lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with dark-light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and			
FEATURES				
SOURCE				

siliques. The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA. "

BASE COUNT 165 a 106 c 135 g 167 t 14 others

Query Match 20.0%; Score 165.4; DB 14; Length 587; Best Local Similarity 74.0%; Pred. No. 4.4e-22; Matches 227; Conservative 0; Mismatches 77; Indels 1; Gaps 1;

19 TCTTCATTGGAGTGAAGAAAGTTCAATGAGCTGATGAAGAGATTTCTTTAGAGAAA 78  
106 TCGCCCAAGAGGTTGAGAGAAAGTGACAGCTGCTGAGAGATGATCTCTTGAGGCTTA 165  
79 TGTATTATAGTATGATGAGAGAAAGGATGATGATGATGATGATGATGATGATGATGAT 138  
166 TGTATTATAGTATGAGAGAAAGGATGATGATGATGATGATGATGATGATGATGATGAT 225  
139 AGATGTCGAGAAAGTTGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 198  
226 CGATGAGAGAAAGTTGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 285  
199 GGTGACTTTGAGCAAGATGAAGTGAATCTATTGAGGCTTCAATAGCTCTTGAAGCAAC 258  
286 GGAAGACTTGAAGCAATGATGAAGTGAATCTATTGAGGCTTCAATAGCTCTTGAAGCAAT 345  
259 A-GATGTCGACTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 317  
346 AGGCTGCTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 405

RESULT 13  
BG881996 397 bp mRNA linear EST 29-NOV-2001  
LOCUS BG881996  
DEFINITION BG881996  
ID: Gm-cl065-3356 5' similar to TR:Q40920 Q40920 MYB-LIKE  
TRANSCRIPTIONAL FACTOR MBFL. ; mRNA sequence.

ACCESSION BG881996  
VERSION BG881996.1 GI:14259088  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 397)  
AUTHORS Shoemaker,R., Kaim,P., Vodkin,L., Expelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schuck,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project

TITLE Unpublished  
JOURNAL Contact: Shoemaker R/Public Soybean EST Project  
COMMENT Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: east@wustl.edu  
This clone is available through: Reggen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800) 533-4363 or contact via email: ccu@reggen.com  
Insert Length: 1212 Std Error: 0.00  
High quality sequence stop: 373.  
Location/Qualifiers  
1..397  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-3356"

## FEATURES

1..397  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-3356"

/tissue\_type="germinating shoots"

/lab\_host="DH10B"

/clone\_id="Gm-cl065"

/note=Vector: pBluescript II SK+, Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated germinating shoots of the cultivar Williams. The seeds were allowed to germinate for 24 hours prior to being cold stressed for 2 days at 4C. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 156 a 72 c 92 g 77 t

Query Match 19.1%; Score 157.4; DB 12; Length 397; Best Local Similarity 66.2%; Pred. No. 1.5e-20; Matches 227; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

9 TACATCATGCTCTTCAATGGAGTGAAGAAAGTTCAATGAGCTGATGAAGAGATTTCT 68  
19 TCTTGTGTTTCAAGAGTGGGGTTGCAAGAGTTCATGATCTCTTAAGAGATGCATT 78  
69 TCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 128  
79 GCTTACCAAGATATATCAAGCTCATGAGAGAGCCATGGAATCACTACCCAAAAAGC 138  
129 TGGTCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 188  
139 AGGCTCTTGAAGTGGAGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 198  
189 TATCAAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 248  
199 CATTAAGAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 258  
249 CTTAGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 308  
259 TTGGGAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 318  
309 GAAAGCTATGGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 351  
319 AAGAGACTGAGAACCCATCTTAAGCAAAAAGCTGAATAAT 361

RESULT 14  
BU830456 781 bp mRNA linear EST 15-OCT-2002  
LOCUS BU830456  
DEFINITION BU830456  
ID: Gm-cl065-3356 5' similar to TR:Q40920 Q40920 MYB-LIKE  
TRANSCRIPTIONAL FACTOR MBFL. ; mRNA sequence.

ACCESSION BU830456  
VERSION BU830456.1 GI:24007683  
KEYWORDS EST.  
SOURCE Populus tremula x Populus tremuloides  
ORGANISM Populus tremula x Populus tremuloides

REFERENCE 1 (bases 1 to 781)  
AUTHORS Uneberg,P., Bhalerao,R.R., Jansson,S. and Sterky,F.  
TITLE The poplar tree transcriptome: Analysis of expressed sequence tags from multiple libraries

JOURNAL Unpublished  
COMMENT Contact: BHALERAO RUPALI R.  
Umea Plant Science Center  
Department of Plant Physiology  
University of Umea, 901 87 Umea, Sweden  
Tel: +46 90 786 5279  
Fax: +46 90 786 6676







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 08:33:17 ; Search time 2209 Seconds

(without alignments)  
15278.607 Million cell updates/sec

Title: US-10-033-190-1

Sequence: 1 atgaacagctacatcatgtc.....tatgaatcctacttgattaa 825

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapect 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl:\*

1: gb\_da:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_da:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_mus:\*

33: em\_htg\_other:\*

34: em\_htg\_pln:\*

35: em\_htg\_rnd:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	3.2	865	8 AF146703	AF146703 Pecunia i
2	26	3.2	866	8 AF146704	AF146704 Pecunia i
3	26	3.2	868	8 AF146706	AF146706 Pecunia x
4	26	3.2	1034	8 AF146702	AF146702 Pecunia x
5	23	2.8	1339	8 AF371977	AF371977 Arabidops
6	23	2.8	59372	8 AB023039	AB023039 Arabidops
7	23	2.8	23747	2 AC116259	AC116259 Rattus no
8	22	2.7	865	8 AF146709	AF146709 Pecunia a
9	22	2.7	868	8 AF146707	AF146707 Pecunia x
10	22	2.7	909	8 AF146705	AF146705 Pecunia x
11	22	2.7	194553	9 AC097536	AC097536 Homo sapi
12	22	2.7	219481	2 AC130926	AC130926 Rattus no
13	22	2.7	220494	2 AC103077	AC103077 Rattus no
14	21	2.5	10913	1 U32791	U32791 Haemophilus
15	21	2.5	26849	1 AE001227	AE001227 Treponema
16	21	2.5	88119	9 AL450325	AL450325 Human DNA
17	21	2.5	110000	6 AR274513_11	AR274513_11 o
18	21	2.5	156730	2 AC025063	AC025063 Homo sapi
19	21	2.5	157268	2 BX255922	BX255922 Danio rer
20	21	2.5	175067	2 AC040898	AC040898 Homo sapi
21	21	2.5	197337	2 BX294661	BX294661 Danio rer
22	21	2.5	199735	2 AC125328	AC125328 Mus muscu
23	21	2.5	213888	2 AC122506	AC122506 Mus muscu
24	21	2.5	225346	2 AC133665	AC133665 Rattus no
25	21	2.5	244067	2 AC110311	AC110311 Rattus no
26	21	2.5	253924	3 AE014822	AE014822 Plasmodiu
27	20	2.4	135	8 AT297048	AT297048 A. thaliana
28	20	2.4	740	6 AX651298	AX651298 Sequence
29	20	2.4	741	6 AX505426	AX505426 Sequence
30	20	2.4	858	8 AF146708	AF146708 Pecunia a
31	20	2.4	1560	5 AF062086	AF062086 Salmo sal
32	20	2.4	1569	3 AK115115	AK115115 Ciona int
33	20	2.4	1789	3 AK113924	AK113924 Ciona int
34	20	2.4	3234	8 NTAI31837	NTAI31837 Nicotiana
35	20	2.4	11596	1 AE010582	AE010582 Fusobacte
36	20	2.4	14799	1 AE001032	AE001032 Archaeogl
37	20	2.4	41061	3 AC006604	AC006604 Caenorhab
38	20	2.4	57348	2 AC101362	AC101362 Mus muscu
39	20	2.4	60108	2 AC090854	AC090854 Homo sapi
40	20	2.4	77830	10 AL672057	AL672057 Mouse DNA
41	20	2.4	87714	9 HS968D22	HS968D22 Human DNA
42	20	2.4	122000	9 AC003093	AC003093 Human BAC
43	20	2.4	126181	9 AC074131	AC074131 Homo sapi
44	20	2.4	128294	2 AC136014	AC136014 Rattus no
45	20	2.4	129350	9 AC068802	AC068802 Homo sapi
46	20	2.4	140069	10 AL355005	AL355005 Mouse DNA
47	20	2.4	141439	10 AC105403	AC105403 Mus muscu
48	20	2.4	147986	2 AC102936	AC102936 Mus muscu
49	20	2.4	149854	10 AC117225	AC117225 Mus muscu
50	20	2.4	152699	2 AP005125	AP005125 Oryza sat
51	20	2.4	155007	2 AC084047	AC084047 Trypanoso
52	20	2.4	161868	2 AL330996	AL330996 Homo sapi
53	20	2.4	166257	2 AC084148	AC084148 Homo sapi
54	20	2.4	166639	2 AC091398	AC091398 Bos tauru
55	20	2.4	168813	9 AP000555	AP000555 Homo sapi
56	20	2.4	176517	2 BX276102	BX276102 Danio rer
57	20	2.4	179613	2 AC133503	AC133503 Mus muscu
58	20	2.4	180158	5 AL929114	AL929114 Zebrafish
59	20	2.4	180998	2 AC136510	AC136510 Pan trogl
60	20	2.4	184099	2 BX530096	BX530096 Mus muscu
61	20	2.4	184991	2 AC108999	AC108999 Rattus no
62	20	2.4	200147	2 AC022774	AC022774 Mus muscu
63	20	2.4	202478	9 AC006560	AC006560 Homo sapi
64	20	2.4	214099	2 AC107711	AC107711 Mus muscu
65	20	2.4	220894	10 AL732410	AL732410 Mouse DNA

Pred. No. is the number of results predicted by chance to have a



C 212	19	2.3 167083	10	AC117233	AC117233 Mus muscu	285	19	2.3 268586	2	AC140326	AC140326 Mus muscu	
C 213	19	2.3 167249	2	AC130478	AC130478 Mus muscu	C 286	19	2.3 270844	2	AC097769	AC097769 Rattus no	
C 214	19	2.3 167533	2	AP004806	AP004806 Oryza sat	C 287	19	2.3 271271	2	AC108635	AC108635 Rattus no	
C 215	19	2.3 167847	2	AC134452	AC134452 Mus muscu	C 288	19	2.3 287291	2	AC115384	AC115384 Rattus no	
C 216	19	2.3 167917	2	AC026102	AC026102 Homo sapi	C 289	19	2.3 289708	2	AC102103	AC102103 Mus muscu	
C 217	19	2.3 168132	9	AC087590	AC087590 Homo sapi	C 290	19	2.3 290029	1	AE016978	AE016978 Shigella	
C 218	19	2.3 169501	2	AC105436	AC105436 Sub scroft	C 291	19	2.3 292715	1	AC098382	AC098382 Rattus no	
C 219	19	2.3 169579	2	AC023885	AC023885 Homo sapi	C 292	19	2.3 295741	1	AP002551	AP002551 Escherich	
C 220	19	2.3 169848	10	AL807753	AL807753 Mouse DNA	C 293	19	2.3 298037	3	AE003674	AE003674 Drosophila	
C 221	19	2.3 169857	2	AC128637	AC128637 Rattus no	C 294	19	2.3 300337	1	AE016927	AE016927 Bacteroid	
C 222	19	2.3 191176	10	AC121817	AC121817 Mus muscu	C 295	19	2.3 300437	2	AC105717	AC105717 Rattus no	
C 223	19	2.3 191318	9	AL732578	AL732578 Human DNA	C 296	19	2.3 301629	2	AC103342	AC103342 Rattus no	
C 224	19	2.3 191523	2	AL928666	AL928666 Mus muscu	C 297	19	2.3 301700	1	AP005948	AP005948 Bradyrhiz	
C 225	19	2.3 191692	2	AC023834	AC023834 Mus muscu	C 298	19	2.3 309805	2	AC026340	AC026340 Homo sapi	
C 226	19	2.3 192676	10	AL671706	AL671706 Mouse DNA	C 299	19	2.3 335638	2	AC135235	AC135235 Mus muscu	
C 227	19	2.3 192861	8	ATCHRIV4	AT161584 Arabidops	C 300	18	2.2 135	8	AT295741	AT295741	
C 228	19	2.3 192867	9	AL358815	AL358815 Human DNA	C 301	18	2.2 179	6	AX645533	AX645533 Sequence	
C 229	19	2.3 193479	2	AC110559	AC110559 Mus muscu	C 302	18	2.2 179	6	AX676694	AX676694 Sequence	
C 230	19	2.3 195925	10	AC123852	AC123852 Mus muscu	C 303	18	2.2 160	8	AB073014	AB073014 Vitis lab	
C 231	19	2.3 196519	2	AC127834	AC127834 Rattus no	C 304	18	2.2 223	11	CF411337	CF411337 Canis fam	
C 232	19	2.3 198602	2	AC129307	AC129307 Mus muscu	C 305	18	2.2 298	10	STDIPSP10	STDIPSP10	
C 233	19	2.3 202764	2	AC114673	AC114673 Mus muscu	C 306	18	2.2 311	3	AF472730	AF472730 Bacillus	
C 234	19	2.3 203414	10	AL845296	AL845296 Mouse DNA	C 307	18	2.2 315	3	AF473162	AF473162 Bacillus	
C 235	19	2.3 204167	2	AC112409	AC112409 Rattus no	C 308	18	2.2 316	3	AF472968	AF472968 Bacillus	
C 236	19	2.3 206254	9	AC093283	AC093283 Homo sapi	C 309	18	2.2 316	3	AF473154	AF473154 Bacillus	
C 237	19	2.3 207842	2	AC010276	AC010276 Homo sapi	C 310	18	2.2 360	6	AX367040	AX367040 Sequence	
C 238	19	2.3 210322	10	AL808131	AL808131 Mouse DNA	C 311	18	2.2 380	11	G35245	G35245 crf1e Crypt	
C 239	19	2.3 211349	10	AL627102	AL627102 Mouse DNA	C 312	18	2.2 459	6	AX606913	AX606913 Sequence	
C 240	19	2.3 211544	9	AC025165	AC025165 Homo sapi	C 313	18	2.2 470	11	DM2H128	DM2H128 Drosophila	
C 241	19	2.3 212302	2	AC133945	AC133945 Mus muscu	C 314	18	2.2 484	5	CPTSLRE02	CPTSLRE02 Gallus gall	
C 242	19	2.3 212628	2	AC097283	AC097283 Rattus no	C 315	18	2.2 536	11	BV000254	BV000254 S210P6048	
C 243	19	2.3 212809	2	AC119328	AC119328 Rattus no	C 316	18	2.2 568	11	BV039170	BV039170 S212P6033	
C 244	19	2.3 213363	2	AC098113	AC098113 Rattus no	C 317	18	2.2 579	8	AF502295	AF502295 Cucumis b	
C 245	19	2.3 215041	2	AC119025	AC119025 Rattus no	C 318	18	2.2 652	11	BV075034	BV075034 S212P6053	
C 246	19	2.3 215635	2	AC100381	AC100381 Mus muscu	C 319	18	2.2 681	8	AY198604	AY198604 Arabidops	
C 247	19	2.3 218667	2	AC095722	AC095722 Rattus no	C 320	18	2.2 711	6	AX367026	AX367026 Sequence	
C 248	19	2.3 218699	2	AC117769	AC117769 Mus muscu	C 321	18	2.2 724	11	BV020850	BV020850 S212P6032	
C 249	19	2.3 220190	2	AC108852	AC108852 Mus muscu	C 322	18	2.2 750	6	AX505636	AX505636 Sequence	
C 250	19	2.3 221209	2	AL928661	AL928661 Mus muscu	C 323	18	2.2 750	8	AF344319	AF344319 Arabidops	
C 251	19	2.3 223606	2	AC013479	AC013479 Homo sapi	C 324	18	2.2 750	8	BT006301	BT006301 Arabidops	
C 252	19	2.3 223726	10	AL626807	AL626807 Mouse DNA	C 325	18	2.2 762	6	AX367016	AX367016 Sequence	
C 253	19	2.3 223756	2	AC144415	AC144415 Rattus no	C 326	18	2.2 788	6	AX536958	AX536958 Sequence	
C 254	19	2.3 225187	2	AC118894	AC118894 Rattus no	C 327	18	2.2 888	9	HS425380	HS425380 Homo sapi	
C 255	19	2.3 226724	10	AC125139	AC125139 Mus muscu	C 328	18	2.2 900	8	AF062907	AF062907 Arabidops	
C 256	19	2.3 227340	2	AC095000	AC095000 Rattus no	C 329	18	2.2 955	8	AY086993	AY086993 Arabidops	
C 257	19	2.3 227364	2	AC113732	AC113732 Rattus no	C 330	18	2.2 973	3	AK113792	AK113792 Clona int	
C 258	19	2.3 227427	2	AC096434	AC096434 Rattus no	C 331	18	2.2 973	9	HS461121	HS461121 Homo sapi	
C 259	19	2.3 227743	2	AC121475	AC121475 Rattus no	C 332	18	2.2 992	8	AF325124	AF325124 Arabidops	
C 260	19	2.3 228736	2	AC107161	AC107161 Rattus no	C 333	18	2.2 999	6	AR316360	AR316360 Sequence	
C 261	19	2.3 230221	10	AC123851	AC123851 Mus muscu	C 334	18	2.2 1005	8	BT001182	BT001182 Arabidops	
C 262	19	2.3 231183	2	AC100052	AC100052 Mus muscu	C 335	18	2.2 1018	3	AK113894	AK113894 Clona int	
C 263	19	2.3 232614	2	AC109088	AC109088 Rattus no	C 336	18	2.2 1033	6	AX358283	AX358283 Sequence	
C 264	19	2.3 234867	2	AC097887	AC097887 Rattus no	C 337	18	2.2 1033	8	AF062915	AF062915 Arabidops	
C 265	19	2.3 235711	2	AC110850	AC110850 Rattus no	C 338	18	2.2 1050	8	AF175989	AF175989 Arabidops	
C 266	19	2.3 235937	2	AC131223	AC131223 Rattus no	C 339	18	2.2 1065	8	SCYDR056C	SCYDR056C S.cerevisia	
C 267	19	2.3 237279	2	AC128607	AC128607 Rattus no	C 340	18	2.2 1114	8	BT005642	BT005642 Arabidops	
C 268	19	2.3 237321	2	AC097391	AC097391 Rattus no	C 341	18	2.2 1140	10	CGYMTCTBC	CGYMTCTBC	
C 269	19	2.3 238382	2	AC122085	AC122085 Rattus no	C 342	18	2.2 1148	8	AF386932	AF386932 Arabidops	
C 270	19	2.3 240583	2	AC107518	AC107518 Rattus no	C 343	18	2.2 1160	10	BC024797	BC024797 Mus muscu	
C 271	19	2.3 241170	2	AC111268	AC111268 Rattus no	C 344	18	2.2 1171	9	AF276882	AF276882 Homo sapi	
C 272	19	2.3 241373	2	AC126747	AC126747 Rattus no	C 345	18	2.2 1171	9	AF276882	AF276882 Homo sapi	
C 273	19	2.3 241430	2	AC123274	AC123274 Rattus no	C 346	18	2.2 1224	8	AF386959	AF386959 Arabidops	
C 274	19	2.3 242112	10	AC122923	AC122923 Mus muscu	C 347	18	2.2 1266	6	AX413845	AX413845 Sequence	
C 275	19	2.3 242849	2	AC113640	AC113640 Rattus no	C 348	18	2.2 1266	6	AX413665	AX413665 Sequence	
C 276	19	2.3 243480	2	AC134378	AC134378 Rattus no	C 349	18	2.2 1341	3	AK113459	AK113459 Clona int	
C 277	19	2.3 245450	2	AC125586	AC125586 Rattus no	C 350	18	2.2 1341	3	AK114822	AK114822 Clona int	
C 278	19	2.3 246075	2	AC111469	AC111469 Rattus no	C 351	18	2.2 1376	9	HSJ001383	HSJ001383 Homo sapi	
C 279	19	2.3 246245	2	AC127221	AC127221 Rattus no	C 352	18	2.2 1400	8	AK118728	AK118728 Arabidops	
C 280	19	2.3 248936	2	AC135200	AC135200 Rattus no	C 353	18	2.2 1400	8	AF126399	AF126399 Arabidops	
C 281	19	2.3 250770	2	AC140246	AC140246 Mus muscu	C 354	18	2.2 1416	6	AX367027	AX367027 Sequence	
C 282	19	2.3 252978	2	AC097297	AC097297 Rattus no	C 355	18	2.2 1431	8	BT000009	BT000009 Arabidops	
C 283	19	2.3 262050	1	AP000985	AP000985 Dicotlobu	C 356	18	2.2 1455	6	AX653098	AX653098 Sequence	
C 284	19	2.3 266544	3	AC116956	AC116956 Dicotlobu	C 357	18	2.2 1467	6	AX367018	AX367018 Sequence	

C 358	18	2.2	1486	8	AY080630	Arabidops	431	18	2.2	43768	9	AC004004	AC004004 Homo sapi
C 359	18	2.2	1674	9	HSMPRA464	X56256 Homo sapien	432	18	2.2	46148	3	AC006624	AC006624 Caenorhab
C 360	18	2.2	1732	3	AKI15598	AKI15598 Ciona int	433	18	2.2	47428	6	AX059539	AX059539 Sequence
C 361	18	2.2	1915	9	AF052102	AF052102 Homo sapi	434	18	2.2	50237	1	EFCA7501	EFCA7501 X92945 Enterococc
C 362	18	2.2	2027	9	AK056199	AK056199 Homo sapi	435	18	2.2	51553	1	AL161780	AL161780 Human DNA
C 363	18	2.2	2145	8	BT001087	BT001087 Arabidops	436	18	2.2	52277	2	AC023540	AC023540 Homo sapi
C 364	18	2.2	2179	8	AY062696	AY062696 Arabidops	437	18	2.2	54351	2	AC137683	AC137683 Cicer ari
C 365	18	2.2	2185	8	AK055664	AK055664 Homo sapi	438	18	2.2	57286	2	AC117683	AC117683 Mus muscu
C 366	18	2.2	2208	6	AX583874	AX583874 Sequence	439	18	2.2	57296	2	AC111159	AC111159 Homo sapi
C 367	18	2.2	2256	6	BD157036	BD157036 Primer fo	440	18	2.2	57655	2	AC100690	AC100690 Mus muscu
C 368	18	2.2	2256	9	AK002073	AK002073 Homo sapi	441	18	2.2	59007	2	AL353757	AL353757 Human DNA
C 369	18	2.2	2337	6	AX540655	AX540655 Sequence	442	18	2.2	63742	2	AC125434	AC125434 Homo sapi
C 370	18	2.2	2397	8	AY057503	AY057503 Arabidops	443	18	2.2	63750	8	AP004936	AP004936 Lotus jap
C 371	18	2.2	2546	8	AY128339	AY128339 Arabidops	444	18	2.2	64016	2	AC101180	AC101180 Mus muscu
C 372	18	2.2	2621	8	BT005783	BT005783 Arabidops	445	18	2.2	64281	9	AP002827	AP002827 Homo sapi
C 373	18	2.2	2783	9	BC007418	BC007418 Homo sapi	446	18	2.2	65723	2	AC130702	AC130702 Homo sapi
C 374	18	2.2	2783	9	BC007447	BC007447 Homo sapi	447	18	2.2	66301	2	AC090550	AC090550 Homo sapi
C 375	18	2.2	2787	6	AX489276	AX489276 Sequence	448	18	2.2	66539	9	HSJ583K4	HSJ583K4 Homo sapi
C 376	18	2.2	3018	3	DMWGR	DMWGR	449	18	2.2	66542	8	TT10118	TT10118 Homo sapi
C 377	18	2.2	3188	9	AK057643	AK057643 Homo sapi	450	18	2.2	66542	8	AP003707	AP003707 Oryza sat
C 378	18	2.2	3188	9	AK057643	AK057643 Homo sapi	451	18	2.2	66542	8	AP003707	AP003707 Oryza sat
C 379	18	2.2	3767	6	AR127480	AR127480 Sequence	452	18	2.2	66776	9	AC092848	AC092848 Homo sapi
C 380	18	2.2	3767	6	AR127481	AR127481 Sequence	453	18	2.2	67359	3	AC092848	AC092848 Homo sapi
C 381	18	2.2	3767	6	AR173301	AR173301 Sequence	454	18	2.2	68843	9	AL450427	AL450427 Human DNA
C 382	18	2.2	3767	6	AR173301	AR173301 Sequence	455	18	2.2	69216	9	AL590230	AL590230 Human DNA
C 383	18	2.2	3933	9	AF084530	AF084530 Homo sapi	456	18	2.2	69475	2	AC101417	AC101417 Mus muscu
C 384	18	2.2	3933	9	AF202144	AF202144 Homo sapi	457	18	2.2	69644	2	AC133771	AC133771 Homo sapi
C 385	18	2.2	3972	9	AF202145	AF202145 Homo sapi	458	18	2.2	70581	9	AL353575	AL353575 Human DNA
C 386	18	2.2	4261	1	AF538722	AF538722 Streptococ	459	18	2.2	70824	2	AC118661	AC118661 Homo sapi
C 387	18	2.2	4402	8	AF245482	AF245482 Oryza sat	460	18	2.2	71688	2	AC134691	AC134691 Homo sapi
C 388	18	2.2	4421	9	AF347478	AF347478 Homo sapi	461	18	2.2	71912	9	AC025443	AC025443 Homo sapi
C 389	18	2.2	5002	14	AF375296	AF375296 Caspalia	462	18	2.2	72453	2	AC013319	AC013319 Homo sapi
C 390	18	2.2	5540	8	BC048423	BC048423 Homo sapi	463	18	2.2	73703	10	AL805916	AL805916 Mouse DNA
C 391	18	2.2	6255	8	AB0057788	AB0057788 Aspergill	464	18	2.2	74237	9	AC098809	AC098809 Papio anu
C 392	18	2.2	6811	1	BSRRALMP	BSRRALMP	465	18	2.2	74571	6	AL449106	AL449106 Human DNA
C 393	18	2.2	7290	10	MMU292467	MMU292467 Mus muscu	466	18	2.2	75248	6	AX602204	AX602204 Sequence
C 394	18	2.2	8207	10	MMU292467	MMU292467 Mus muscu	467	18	2.2	75468	2	AC094704	AC094704 Sequence
C 395	18	2.2	9603	9	AC110793	AC110793 Homo sapi	468	18	2.2	76816	8	AB020749	AB020749 Arabidops
C 396	18	2.2	10237	1	AE006438	AE006438 Lactococc	469	18	2.2	77088	8	AL592296	AL592296 Human DNA
C 397	18	2.2	10396	1	AE010967	AE010967 Methanosa	470	18	2.2	77129	8	AB023033	AB023033 Arabidops
C 398	18	2.2	10506	1	AE011645	AE011645 Xanthomon	471	18	2.2	77439	9	HS127F18	HS127F18 Human DNA
C 399	18	2.2	10811	1	AE011764	AE011764 Xanthomon	472	18	2.2	78735	9	AC006545	AC006545 Homo sapi
C 400	18	2.2	10900	1	AE002310	AE002310 Chlamydia	473	18	2.2	78805	8	OSJUN0123	OSJUN0123 Homo sapi
C 401	18	2.2	11276	2	CNS0984Y	BNX29433 Oryza sat	474	18	2.2	78867	2	AC025322	AC025322 Homo sapi
C 402	18	2.2	15561	2	AC018211	AC018211 Thermocog	475	18	2.2	79128	9	AC006546	AC006546 Homo sapi
C 403	18	2.2	15672	9	AC096584	AC096584 Homo sapi	476	18	2.2	79643	2	AC115790	AC115790 Homo sapi
C 404	18	2.2	15951	6	AX281470	AX281470 Sequence	477	18	2.2	79677	8	ATT22C14	ATT22C14 Arabidops
C 405	18	2.2	15951	6	AX344683	AX344683 Sequence	478	18	2.2	79995	2	AC022231	AC022231 Mus muscu
C 406	18	2.2	15951	6	AX344683	AX344683 Sequence	479	18	2.2	80019	8	ATF25E4	ATF25E4 Arabidops
C 407	18	2.2	16662	1	AE002138	AE002138 Ureaplasma	480	18	2.2	80074	9	AC008389	AC008389 Homo sapi
C 408	18	2.2	16931	1	AE001716	AE001716 Thermocog	481	18	2.2	81172	2	AC036240	AC036240 Homo sapi
C 409	18	2.2	17674	6	AX3446247	AX3446247 Sequence	482	18	2.2	81835	8	ATFCA9	ATFCA9 Arabidops
C 410	18	2.2	20315	1	AE014280	AE014280 Streptococ	483	18	2.2	82074	10	EX321909	EX321909 Mouse DNA
C 411	18	2.2	21229	2	AC015141	AC015141 Drosophill	484	18	2.2	83233	8	AB022219	AB022219 Arabidops
C 412	18	2.2	21343	8	SPBC17A3	AL109652 S.pombe	485	18	2.2	84216	9	AL157718	AL157718 Human DNA
C 413	18	2.2	23343	9	AL591396	AL591396 Human DNA	486	18	2.2	84876	9	BX284698	BX284698 Human DNA
C 414	18	2.2	23709	8	AP004980	AP004980 Lotus jap	487	18	2.2	84881	9	AL390726	AL390726 Human DNA
C 415	18	2.2	27387	2	AC017975	AC017975 Drosophill	488	18	2.2	85702	8	AC020665	AC020665 Arabidops
C 416	18	2.2	28789	9	HS032169	HS032169 Human pro-a	489	18	2.2	85984	2	AC122711	AC122711 Homo sapi
C 417	18	2.2	29244	8	AC124834	AC124834 Homo sapi	490	18	2.2	86050	8	ATF7A7	ATF7A7 Homo sapi
C 418	18	2.2	30017	8	SCCHIYDNA	X84162 S.cerevisia	491	18	2.2	86690	8	ATF7A7	ATF7A7 Homo sapi
C 419	18	2.2	30866	8	AL590393	AL590393 Human DNA	492	18	2.2	86897	9	AC0091843	AC0091843 Homo sapi
C 420	18	2.2	37017	3	U12966	U12966 Caenorhabdi	493	18	2.2	88460	2	AC140721	AC140721 Homo sapi
C 421	18	2.2	37187	8	AB004537	AB004537 Schizosac	494	18	2.2	88460	2	AC140721	AC140721 Homo sapi
C 422	18	2.2	41099	8	CEB0285	CEB0285	495	18	2.2	88987	2	AC004541	AC004541 Homo sapi
C 423	18	2.2	41397	3	U64599	U64599 Caenorhabdi	496	18	2.2	89050	1	BSZ75208	BSZ75208 Homo sapi
C 424	18	2.2	41984	3	CEB0285	CEB0285	497	18	2.2	89370	8	ATT9L3	ATT9L3 Homo sapi
C 425	18	2.2	42112	3	CEB0285	CEB0285	498	18	2.2	90435	2	AC138794	AC138794 Homo sapi
C 426	18	2.2	42660	3	CEB0285	CEB0285	499	18	2.2	90509	8	ATT13K14	ATT13K14 Homo sapi
C 427	18	2.2	43159	8	AF193903	AF193903 Caeteria	500	18	2.2	91570	8	ATT13K14	ATT13K14 Homo sapi
C 428	18	2.2	43178	9	HSU73465	HSU73465 Human Xq28	501	18	2.2	91789	3	AC010839	AC010839 Homo sapi
C 429	18	2.2	43658	2	AC144435	AC144435 Homo sapi	502	18	2.2	91822	9	AL160268	AL160268 Homo sapi
C 430	18	2.2	43658	2	AC144435	AC144435 Homo sapi	503	18	2.2	91822	9	AL160268	AL160268 Homo sapi



650	18	2.2	148386	9	AC104391	Homo sapi	c 723	18	2.2	160091	2	AC025686	AC025686 Homo sapi
c 651	18	2.2	148875	9	AL353736	Human DNA	724	18	2.2	160243	2	AC027223	AC027223 Homo sapi
652	18	2.2	148917	9	AC008942	Homo sapi	725	18	2.2	160404	2	AL159152	AL159152 Human DNA
c 653	18	2.2	149252	9	AC066580	Homo sapi	c 726	18	2.2	160408	2	AC012196	AC012196 Homo sapi
c 654	18	2.2	149303	8	AC087412	Oryza sat	c 727	18	2.2	160724	5	BLX005101	BLX005101 Zedraflish
655	18	2.2	149539	8	AC136284	Genomic s	c 728	18	2.2	160942	9	AL1390966	AL1390966 Human DNA
c 656	18	2.2	149560	2	BX255929	Rattus rer	c 729	18	2.2	160982	2	AC128945	AC128945 Rattus no
657	18	2.2	149848	2	AC119516	Danio rer	c 730	18	2.2	161040	2	AL1359894	AL1359894 Human DNA
658	18	2.2	150271	10	AC125309	Mus muscu	c 731	18	2.2	161094	2	AC012299	AC012299 Homo sapi
659	18	2.2	150489	2	AC069491	Homo sapi	c 732	18	2.2	161109	2	AC093594	AC093594 Homo sapi
c 660	18	2.2	150496	8	AC131968	Oryza sat	c 733	18	2.2	161198	9	AF274854	AF274854 Homo sapi
c 661	18	2.2	150615	2	AC101984	Mus muscu	c 734	18	2.2	161218	9	AC011501	AC011501 Homo sapi
662	18	2.2	150817	8	AP003223	Oryza sat	c 735	18	2.2	161366	2	AC116248	AC116248 Rattus no
663	18	2.2	150884	2	AC009382	Drosophil	736	18	2.2	161375	9	AC068715	AC068715 Homo sapi
c 664	18	2.2	150938	2	AC025293	Homo sapi	c 737	18	2.2	161452	9	AC020925	AC020925 Homo sapi
c 665	18	2.2	151242	2	AC144589	Homo sapi	c 738	18	2.2	161619	2	AC139490	AC139490 Homo sapi
c 666	18	2.2	151751	10	BX088539	Mouse DNA	c 739	18	2.2	161740	2	AC084778	AC084778 Homo sapi
c 667	18	2.2	152132	2	AC011592	Homo sapi	c 740	18	2.2	161845	2	BX005014	BX005014 Danio rer
668	18	2.2	152433	2	AC036127	Homo sapi	c 741	18	2.2	161893	9	AC137781	AC137781 Homo sapi
669	18	2.2	152619	2	AC107977	Homo sapi	c 742	18	2.2	162206	9	AC108706	AC108706 Homo sapi
670	18	2.2	152757	10	AC121515	Rattus no	743-	18	2.2	162384	2	AC034236	AC034236 Homo sapi
671	18	2.2	152865	2	AL359537	Homo sapi	c 744	18	2.2	162404	2	AC102849	AC102849 Mus muscu
672	18	2.2	153065	2	AP005094	Oryza sat	c 745	18	2.2	162474	9	AC022706	AC022706 Homo sapi
673	18	2.2	153190	2	AC026110	Homo sapi	746	18	2.2	162695	2	AC032746	AC032746 Homo sapi
c 674	18	2.2	153999	9	AC131213	Homo sapi	747	18	2.2	162803	2	AC032000	AC032000 Homo sapi
675	18	2.2	154216	2	AC120503	Didelphis	c 748	18	2.2	162955	2	AC012527	AC012527 Homo sapi
c 676	18	2.2	154265	2	AP006446	Oryza sat	c 749	18	2.2	163043	2	AC124654	AC124654 Homo sapi
c 677	18	2.2	154262	2	AC127629	Rattus no	750	18	2.2	163209	2	AC007949	AC007949 Homo sapi
c 678	18	2.2	154495	2	AC127223	Mus muscu	c 751	18	2.2	163291	9	AC022459	AC022459 Homo sapi
c 679	18	2.2	154706	2	AC007964	Homo sapi	c 752	18	2.2	163390	2	AC023369	AC023369 Homo sapi
c 680	18	2.2	154904	2	AL162272	Human DNA	753	18	2.2	163331	9	AC092802	AC092802 Homo sapi
681	18	2.2	155231	2	AC124405	Mus muscu	c 754	18	2.2	163408	9	AC079762	AC079762 Homo sapi
682	18	2.2	155637	10	AC111012	Mus muscu	c 755	18	2.2	163719	9	AC009147	AC009147 Homo sapi
c 683	18	2.2	156246	5	BX005283	Zedraflish	756	18	2.2	163795	2	EX005192	EX005192 Human DNA
c 684	18	2.2	156371	2	AC139332	Mus muscu	757	18	2.2	163865	2	AC102221	AC102221 Mus muscu
c 685	18	2.2	156393	2	AC098652	Homo sapi	758	18	2.2	163865	2	AC138835	AC138835 Homo sapi
c 686	18	2.2	156441	9	AP004074	Homo sapi	c 759	18	2.2	164239	9	AC027541	AC027541 Homo sapi
687	18	2.2	156441	9	AC091869	Homo sapi	c 760	18	2.2	164500	2	AL1359698	AL1359698 Homo sapi
c 688	18	2.2	156572	10	AL845501	Mouse DNA	c 761	18	2.2	164728	2	AC026101	AC026101 Homo sapi
c 689	18	2.2	156608	2	AC015512	Homo sapi	c 762	18	2.2	165178	2	AC127080	AC127080 Rattus no
c 690	18	2.2	156876	2	AP001353	Homo sapi	763	18	2.2	165372	2	EX005153	EX005153 Danio rer
c 691	18	2.2	156902	2	AC073630	Homo sapi	c 764	18	2.2	165790	2	AC009809	AC009809 Homo sapi
c 692	18	2.2	157010	5	BX004854	Zedraflish	c 765	18	2.2	165882	2	AL1354852	AL1354852 Human DNA
c 693	18	2.2	157189	3	AC008307	Drosophil	766	18	2.2	165990	2	AC093287	AC093287 Homo sapi
c 694	18	2.2	157195	2	AC122270	Mus muscu	c 767	18	2.2	166009	2	AC111145	AC111145 Mus muscu
c 695	18	2.2	157198	2	AC024937	Homo sapi	c 768	18	2.2	166160	3	AC011072	AC011072 Drosophil
c 696	18	2.2	157198	2	AC012103	Homo sapi	c 769	18	2.2	166176	2	AC021219	AC021219 Homo sapi
c 697	18	2.2	157382	2	AC009279	Homo sapi	c 770	18	2.2	166310	2	AC114237	AC114237 Homo sapi
c 698	18	2.2	157390	2	AP004283	Homo sapi	771	18	2.2	166455	10	AC124132	AC124132 Rattus no
699	18	2.2	157463	2	BX511126	Danio rer	c 772	18	2.2	166731	9	AC093910	AC093910 Homo sapi
700	18	2.2	157557	2	AC098808	Papio anu	c 773	18	2.2	166793	2	AC090401	AC090401 Homo sapi
c 701	18	2.2	157574	9	AL592293	Human DNA	c 774	18	2.2	166859	2	EX511001	EX511001 Danio rer
702	18	2.2	157713	10	AC121936	Mus muscu	c 775	18	2.2	166928	2	AC055861	AC055861 Homo sapi
c 703	18	2.2	157907	2	BX510302	Danio rer	776	18	2.2	166928	2	AC022631	AC022631 Homo sapi
c 704	18	2.2	157965	2	AC117958	Homo sapi	777	18	2.2	167190	9	AC023301	AC023301 Homo sapi
705	18	2.2	157968	2	AC109315	Homo sapi	778	18	2.2	167203	8	AP003264	AP003264 Oryza sat
c 706	18	2.2	158033	9	AL645940	Human DNA	779	18	2.2	167283	9	AL133479	AL133479 Human DNA
c 707	18	2.2	158091	9	AC140171	Homo sapi	780	18	2.2	167497	2	AL133479	AL133479 Human DNA
c 708	18	2.2	158146	9	AC005076	Homo sapi	c 781	18	2.2	167851	3	AC009381	AC009381 Drosophil
c 709	18	2.2	158424	2	CNS01R13	Homo sapi	c 782	18	2.2	167851	9	AC118660	AC118660 Homo sapi
c 710	18	2.2	158424	9	AC0096878	Human chr	783	18	2.2	167869	9	AC077342	AC077342 Homo sapi
c 711	18	2.2	158646	9	AC142307	Pan trogl	784	18	2.2	167934	9	AP001034	AP001034 Homo sapi
c 712	18	2.2	158841	9	AC007879	Homo sapi	785	18	2.2	167934	2	AC013304	AC013304 Homo sapi
c 713	18	2.2	158917	2	AC101892	Mus muscu	786	18	2.2	168075	2	AC079805	AC079805 Homo sapi
c 714	18	2.2	159284	2	AC0084369	Homo sapi	787	18	2.2	168185	2	AC009381	AC009381 Homo sapi
c 715	18	2.2	159312	2	AC096878	Pan trogl	788	18	2.2	168243	2	AC011269	AC011269 Homo sapi
c 716	18	2.2	159322	2	AC096631	Homo sapi	789	18	2.2	168384	3	AC011069	AC011069 Homo sapi
c 717	18	2.2	159336	2	AC020584	Homo sapi	c 790	18	2.2	168430	8	AC126221	AC126221 Oryza sat
c 718	18	2.2	159469	2	AC079618	Homo sapi	c 791	18	2.2	168448	9	AC009469	AC009469 Homo sapi
719	18	2.2	159557	2	AC121267	Mus muscu	792	18	2.2	168610	2	AC073503	AC073503 Homo sapi
720	18	2.2	159941	2	AC006678	Caenorhab	c 793	18	2.2	168612	2	AC138831	AC138831 Homo sapi
c 721	18	2.2	160039	9	AC068323	Homo sapi	794	18	2.2	168742	2	AC138640	AC138640 Mus muscu
722	18	2.2	160050	1	AL591974	Listeria	795	18	2.2	168866	2	AC136391	AC136391 Rattus no





942	18	2.2	189121	2	AC102914	AC102914 Mus muscu
C 943	18	2.2	189596	3	AC008313	AC008313 Drosophi
944	18	2.2	189712	2	AC023854	AC023854 Homo sapi
945	18	2.2	189712	2	AC012181	AC012181 Homo sapi
C 946	18	2.2	189963	2	AC023376	AC023376 Homo sapi
947	18	2.2	189963	2	AC018822	AC018822 Homo sapi
948	18	2.2	190177	2	AC009959	AC009959 Homo sapi
949	18	2.2	190388	2	AC141348	AC141348 Rattus no
950	18	2.2	190470	2	AC132685	AC132685 Mus muscu
951	18	2.2	190786	2	AC137521	AC137521 Medicago
C 952	18	2.2	190861	2	AC083855	AC083855 Homo sapi
C 953	18	2.2	190927	2	AC141748	AC141748 Apis mell
954	18	2.2	191028	2	AC140168	AC140168 Homo sapi
955	18	2.2	191132	2	AC010282	AC010282 Homo sapi
956	18	2.2	191161	2	AC018603	AC018603 Homo sapi
957	18	2.2	191544	2	AC138855	AC138855 Homo sapi
C 958	18	2.2	191810	2	AC137503	AC137503 Homo sapi
C 959	18	2.2	191835	2	AC110195	AC110195 Mus muscu
960	18	2.2	192116	2	AC109158	AC109158 Mus muscu
C 961	18	2.2	192243	2	AC127939	AC127939 Rattus no
C 962	18	2.2	192925	2	AC109540	AC109540 Rattus no
963	18	2.2	192963	2	AL354763	AL354763 Human DNA
964	18	2.2	193100	2	AC016729	AC016729 Homo sapi
965	18	2.2	193161	2	AC009561	AC009561 Homo sapi
966	18	2.2	193364	2	AC087350	AC087350 Homo sapi
967	18	2.2	193550	8	ATCHRIV58	ATCHRIV58 Arabidops
C 968	18	2.2	193676	2	AC018932	AC018932 Homo sapi
969	18	2.2	194052	2	AC053537	AC053537 Homo sapi
C 970	18	2.2	194149	2	AC100799	AC100799 Homo sapi
971	18	2.2	194170	2	AC138401	AC138401 Mus muscu
C 972	18	2.2	194254	2	AC100811	AC100811 Homo sapi
973	18	2.2	194350	2	AC128084	AC128084 Rattus no
974	18	2.2	194378	2	AL645770	AL645770 Homo sapi
975	18	2.2	194399	2	AP001125	AP001125 Homo sapi
C 976	18	2.2	194943	9	AC105252	AC105252 Homo sapi
C 977	18	2.2	195246	10	AC122266	AC122266 Mus muscu
978	18	2.2	195514	2	AC109217	AC109217 Mus muscu
979	18	2.2	195981	9	AC007652	AC007652 Homo sapi
C 980	18	2.2	196085	2	AC025971	AC025971 Homo sapi
981	18	2.2	196553	2	BX119953	BX119953 Mus muscu
C 982	18	2.2	196721	9	AC019179	AC019179 Homo sapi
983	18	2.2	196852	9	AC084866	AC084866 Homo sapi
C 984	18	2.2	196978	2	AC015872	AC015872 Homo sapi
C 985	18	2.2	197363	9	AC114975	AC114975 Homo sapi
C 986	18	2.2	197568	8	ATCHRIV53	ATCHRIV53 Arabidops
C 987	18	2.2	197775	8	AC012366	AC012366 Homo sapi
C 988	18	2.2	198067	8	ATCHRIV47	ATCHRIV47 Arabidops
C 989	18	2.2	198426	10	AL929433	AL929433 Mouse DNA
990	18	2.2	198563	8	ATCHRIV14	ATCHRIV14 Arabidops
C 991	18	2.2	198587	2	AC005034	AC005034 Homo sapi
992	18	2.2	198611	2	AC108802	AC108802 Homo sapi
993	18	2.2	198788	8	ATCHRIV32	ATCHRIV32 Arabidops
C 994	18	2.2	199090	10	AC116572	AC116572 Mus muscu
C 995	18	2.2	199386	2	AC117440	AC117440 Homo sapi
996	18	2.2	199558	2	AP001807	AP001807 Homo sapi
C 997	18	2.2	199636	2	AP001126	AP001126 Homo sapi
C 998	18	2.2	199814	9	AC023480	AC023480 Homo sapi
999	18	2.2	199934	9	AC019086	AC019086 Homo sapi
1000	18	2.2	199954	9	AC009498	AC009498 Homo sapi

## ALIGNMENTS

RESULT 1  
LOCUS AFI146703  
DEFINITION Petunia integrifolia An2 protein (an2) mRNA, an2-S9 allele,  
complete cds.  
ACCESSION AFI146703.1 GI:7673085  
VERSION AFI146703.1  
KEYWORDS  
SOURCE Petunia integrifolia

ORGANISM  
Petunia integrifolia  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; lamiales; Solanales; Solanaceae; Petunia.  
REFERENCE  
AUTHORS  
Mol, J., and Koes, R.  
TITLE  
Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color  
JOURNAL  
Plant Cell 11 (8), 1433-1444 (1999)  
MEDLINE  
99380006  
PUBMED  
10449578  
REFERENCE  
2 (bases 1 to 865)  
Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N.,  
Mol, J., and Koes, R.  
AUTHORS  
Mol, J., and Koes, R.  
TITLE  
Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan  
1087, Amsterdam 1081HV, Netherlands  
JOURNAL  
Location/Qualifiers  
FEATURES  
source  
1..865  
/organism="Petunia integrifolia"  
/mol\_type="mRNA"  
/cultivar="S9"  
/db\_xref="taxon:4103"  
/chromosome="6"  
/tissue\_type="petal limb"  
1..865  
/gene="an2"  
/allele="an2-S9"  
12..779  
/gene="an2"  
/note="anthocyanin regulator; myb-related transcription  
factor"

gene  
/codon\_start=1  
/product="An2 protein"  
/protein\_id="AF146728.1"  
/db\_xref="GI:7673086"  
/translation="MSTSNASTSGVKGAWTEEDLLREICRYGEGKMLVPRAG  
LNRGRSCRRLMYRLPHIKGGDFSLDEVDLILRLKILGRMSLGLPRTAND  
VKNYMTDHEKGLIAPDOKSKAMKTTENSIIKPRPRTSRPAMNIVSCMGSG  
CNKQTTIDNEGDTREITIKFGDEKRPRESYIDGLOWMNLANNIEBELVSNPTLL  
HEETAPRVNNESSLTQEGSGSLDFSVDDIDVLLS"  
BASE COUNT  
309 a 133 c 193 g 229 t 1 others  
ORIGIN  
Query Match 3.2%; Score 26; DB 8; Length 865;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 223 GATTCATTTTGGAGGCTTCAAGCT 248  
|||||  
Db 234 GATTCATTTTGGAGGCTTCAAGCT 259  
|||||  
RESULT 2  
LOCUS AFI146704  
DEFINITION Petunia integrifolia An2 protein (an2) mRNA, an2-S6 allele,  
complete cds.  
ACCESSION AFI146704  
VERSION AFI146704.1 GI:7673087  
KEYWORDS  
SOURCE  
ORGANISM  
Petunia integrifolia  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; lamiales; Solanales; Solanaceae; Petunia.  
REFERENCE  
AUTHORS  
Mol, J., and Koes, R.  
TITLE  
Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N.,  
Mol, J., and Koes, R.  
JOURNAL  
Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color  
Plant Cell 11 (8), 1433-1444 (1999)



MEDLINE 99380006  
PUBMED 10449578  
REFERENCE 2 (bases 1 to 866)  
AUTHORS Quattrocchio,F., Wang,J., van der Woude,K., Souer,E., de Vetten,N.,  
Mol,J., and Koes,R.  
TITLE Direct Submission  
JOURNAL Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan  
1087, Amsterdam 1081HV, Netherlands  
location/Qualifiers  
FEATURES  
source 1..866  
/organism="Petunia integrifolia"  
/mol\_type="mRNA"  
/cultiivar="S6"  
/db\_xref="taxon:4103"  
/chromosome="6"  
/tissue\_type="petal limb"  
1..866  
/gene="an2"  
/allele="an2-S6"  
12..779  
/gene="an2"  
/note="anthocyanin regulator; myb-related transcription  
factor"  
/codon\_start=1  
/product="An2 protein"  
/protein\_id="AAF66729.1"  
/db\_xref="GI:7673088"  
/translation="MSTSNASTSGVRKANTGEEEDLLRECIKYGEGKHLVPRAG  
LNRCKRSCRLWNLTPHTRKGFSLDEVDLILRLKLGNRWLSIAGRLPORTAND  
VKNYWNTHLRKLIAPHDQKESKSKS  
CSNKTIDKNEGDTEIIFKSDKCPRESIDGLOMANLANNIEBELVSCNSPTLL  
HEETAPSPVNESSLTQCGSGSLSDPSIDIDIMDLVS"  
BASE COUNT 310 a 134 c 195 g 227 t  
ORIGIN  
Query Match 3.2%; Score 26; DB 8; Length 866;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 223 GATCTCATTTTGAGGCTTCATAAGCT 248  
234 GATCTCATTTTGAGGCTTCATAAGCT 259  
Db 234 GATCTCATTTTGAGGCTTCATAAGCT 259  
RESULT 3  
AF146706 868 bp mRNA linear PLN 01-MAY-2000  
LOCUS Petunia x hybrida An2 truncated protein (an2) mRNA, an2-W22 allele,  
DEFINITION complete cds.  
ACCESSION AF146706  
VERSION AF146706.1 GI:7673091  
KEYWORDS Petunia x hybrida  
SOURCE Petunia x hybrida  
ORGANISM Petunia x hybrida  
REFERENCE 1 (bases 1 to 868)  
AUTHORS Quattrocchio,F., Wang,J., van der Woude,K., Souer,E., de Vetten,N.,  
Mol,J., and Koes,R.  
TITLE Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color  
JOURNAL Plant Cell 11 (8), 1433-1444 (1999)  
MEDLINE 99380006  
PUBMED 10449578  
REFERENCE 2 (bases 1 to 868)  
AUTHORS Quattrocchio,F., Wang,J., van der Woude,K., Souer,E., de Vetten,N.,  
Mol,J., and Koes,R.  
TITLE Direct Submission  
JOURNAL Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan  
1087, Amsterdam 1081HV, Netherlands  
location/Qualifiers  
FEATURES  
source 1..868

/organism="Petunia x hybrida"  
/mol\_type="mRNA"  
/cultiivar="W22"  
/db\_xref="taxon:4102"  
/chromosome="6"  
/tissue\_type="petal limb"  
1..868  
/gene="an2"  
/allele="an2-W22"  
12..395  
/gene="an2"  
/note="anthocyanin regulator; myb-related transcription  
factor"  
/codon\_start=1  
/product="An2 truncated protein"  
/protein\_id="AAF66731.1"  
/db\_xref="GI:7673092"  
/translation="MSTSNASTSGVRKANTGEEEDLLRECIKYGEGKHLVPRAG  
LNRCKRSCRLWNLTPHTRKGFSLDEVDLILRLKLGNRWLSIAGRLPORTAND  
VKNYWNTHLRKLIAPHDQKESKSKS"  
392..395  
/gene="an2"  
/note="causes frameshift leading to truncated protein"  
/replace="nnnn"  
BASE COUNT 310 a 140 c 191 g 227 t  
ORIGIN  
Query Match 3.2%; Score 26; DB 8; Length 868;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 223 GATCTCATTTTGAGGCTTCATAAGCT 248  
234 GATCTCATTTTGAGGCTTCATAAGCT 259  
Db 234 GATCTCATTTTGAGGCTTCATAAGCT 259  
RESULT 4  
AF146702 1034 bp mRNA linear PLN 01-MAY-2000  
LOCUS Petunia x hybrida An2 protein (an2) mRNA, an2-V26 allele, complete  
DEFINITION cds.  
ACCESSION AF146702  
VERSION AF146702.1 GI:7673083  
KEYWORDS Petunia x hybrida  
SOURCE Petunia x hybrida  
ORGANISM Petunia x hybrida  
REFERENCE 1 (bases 1 to 1034)  
AUTHORS Quattrocchio,F., Wang,J., van der Woude,K., Souer,E., de Vetten,N.,  
Mol,J., and Koes,R.  
TITLE Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color  
JOURNAL Plant Cell 11 (8), 1433-1444 (1999)  
MEDLINE 99380006  
PUBMED 10449578  
REFERENCE 2 (bases 1 to 1034)  
AUTHORS Quattrocchio,F., Wang,J., van der Woude,K., Souer,E., de Vetten,N.,  
Mol,J., and Koes,R.  
TITLE Direct Submission  
JOURNAL Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan  
1087, Amsterdam 1081HV, Netherlands  
location/Qualifiers  
FEATURES  
source 1..1034  
/organism="Petunia x hybrida"  
/mol\_type="mRNA"  
/cultiivar="V26"  
/db\_xref="taxon:4102"  
/chromosome="6"  
/tissue\_type="petal limb"  
1..1034  
/gene="an2"

CDS  
/allele="an2-V26"  
100...867  
/gene="an2"  
/note="anthocyanin regulator; myb-related transcription factor"  
/codon\_start=1  
/product="An2 protein"  
/protein\_id="AAF6727.1"  
/db\_xref="GI:7673084"  
/translation="WSTSNASTSGVRKGAATTEEDLLIRECIDKYGKQKMLVYVRG  
LNRCKSCRLRWLNYLRPHIKRGDFSLDEVLLIRLHLGNRSLIAGRLPGTAD  
VKNYWTHLRKLLIAPDKOKESKKAIVKTIENNIIPRPTEPRPMNNFPCMGNS  
CKNKTIDNKEDTEIIKESDEKOKESIDGGLQMMANLANNIEIELVSCNSPTLL  
HEETAPVSNESLITGGGSGLSDFVDDIDIMVLVS"

BASE COUNT 365 a 154 c 216 g 299 t

ORIGIN

Query Match 3.2%; Score 26; DB 8; Length 1034;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 223 GATCTCATTTGAGGCTTCATAGCT 248  
|||||  
322 GATCTCATTTGAGGCTTCATAGCT 347

Db

RESULT 5  
AF371977 1339 bp mRNA linear PLN 02-SEP-2001  
LOCUS Arabidopsis thaliana putative transcription factor (At5g49330)  
DEFINITION AF371977  
VERSION AF371977 GI:15420625  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 1339)  
Stracke,R., Weiber,M. and Weisshaar,B.  
The R2R3-MYB gene family in Arabidopsis thaliana  
Curr. Opin. Plant Biol. 4 (5), 447-456 (2001)  
21481677  
PUBMED 11597504  
2 (bases 1 to 1339)  
Weiber,M., Stracke,R. and Weisshaar,B.  
Direct Submision  
Submitted (18-Apr-2001) Dept. Plant Breeding and Yield Physiology,  
Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg  
10, Koeln D-50829, Germany  
Location/Qualifiers

FEATURES  
source  
1..1339  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/note="ecotype: Columbia"  
1..1339  
/gene="At5g49330"  
1..85  
/gene="At5g49330"  
86..1114  
/gene="At5g49330"  
/note="ATMYB11; R2R3-MYB family member"  
/codon\_start=1  
/product="putative transcription factor"  
/protein\_id="AAK97396.1"  
/db\_xref="GI:15420625"  
/translation="WGRAPCECKGKRGRTAEDEILTYIQNGSGSRSLPKKA  
GLRCKGKCRRLRWLNYLRPHIKRGDFSLDEVLLIRLHLGNRSLIAGRLPGTAD  
VKNYWTHLRKLLIAPDKOKESKKAIVKTIENNIIPRPTEPRPMNNFPCMGNS  
CKNKTIDNKEDTEIIKESDEKOKESIDGGLQMMANLANNIEIELVSCNSPTLL  
SRSWKQKQVTAOCPSQPELESDEGQNGNPFGEIEMLDGELERLLSS  
CWECTSEAVIAGVNDKVCSSGDNSSCVNLPFEEQGSERKIGHVIGITEVDHDMTVE

3' UTR  
1115..1339  
/gene="At5g49330"  
BASE COUNT 440 a 184 c 351 g 364 t

ORIGIN

Query Match 2.8%; Score 23; DB 8; Length 1339;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 253 GGCAACAGATGTCATTATTC 275  
|||||  
341 GGCAACAGATGTCATTATTC 363

Db

RESULT 6  
AB023039 59372 bp DNA linear PLN 27-DEC-2000  
LOCUS Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:MIF21.  
DEFINITION AB023039 BA000015  
ACCESSION AB023039.1 GI:4220638  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (sites)  
Sato,S., Nakamura,Y., Kaneko,T., Katoh,T., Asamizu,E., Kotani,H.  
and Tabata,S.  
Structural analysis of Arabidopsis thaliana chromosome 5. X.  
Sequence features of the regions of 3,076,755 bp covered by sixty  
P1 and TAC clones  
DNA Res. 7 (1), 31-63 (2000)  
20181125  
PUBMED 10718197  
2 (bases 1 to 59372)  
Nakamura,Y.  
Direct Submision  
Submitted (01-FEB-1999) Yasukazu Nakamura, Kazusa DNA Research  
Institute, Department of Plant Gene Research, 1532-3, Yana,  
Kisarazu, Chiba 292-0812, Japan (E-mail:yakam@kazusa.or.jp,  
Tel:81-438-52-3935, Fax:81-438-52-3934)  
Address for correspondence: kaos@kazusa.or.jp  
For the latest information on annotation of this clone, please see  
http://www.kazusa.or.jp/kaos/cgi-bin/sgd\_graph.cgi?c=MIF21  
Genes with similarity to proteins in the databases are described in  
'product' or 'note' qualifiers. Genes that have no significant  
protein similarity are described as 'unknown protein'.  
The software programs used to predict genes include: Grail  
(Informatics Group, Oak Ridge National Laboratory,  
http://compbio.ornl.gov/Grail-1.3/),  
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),  
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of  
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and  
SplicePredictor (Volker Brendel, Stanford University,  
http://gremlim.zool.iastate.edu/cgi-bin/sp.cgi).  
Genes encoding tRNAs are predicted by tRNAscan-SB  
(Sean Eddy, Washington University School of Medicine, St. Louis,  
http://genome.wustl.edu/eddy/tRNAscan-SB/).  
This sequence may not be the entire insert of this clone. It may be  
shorter because we remove overlaps between neighboring submissions.  
The 5' clone is MDN11 and the 3' clone is K23f3.  
Location/Qualifiers

FEATURES  
source  
1..59372  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="MIF21"  
/clone\_1id="Mitsui P1"

exon  
complement(1..30)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=18  
/evidence=not\_experimental  
complement(109..263)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=17  
/evidence=not\_experimental  
complement(587..698)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=16  
/evidence=not\_experimental  
complement(770..845)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=15  
/evidence=not\_experimental  
complement(1223..2061)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=14  
/evidence=not\_experimental  
complement(2161..2266)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=13  
/evidence=not\_experimental  
complement(2831..2952)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=12  
/evidence=not\_experimental  
complement(3033..3412)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=11  
/evidence=not\_experimental  
complement(3561..4266)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=10  
/evidence=not\_experimental  
complement(4367..4453)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=9  
/evidence=not\_experimental  
complement(4910..4963)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=8  
/evidence=not\_experimental  
complement(5062..5250)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=7  
/evidence=not\_experimental  
complement(5444..5503)

exon  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=6  
/evidence=not\_experimental  
complement(5603..5790)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=5  
/evidence=not\_experimental  
complement(5859..5936)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=4  
/evidence=not\_experimental  
complement(6035..6089)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=3  
/evidence=not\_experimental  
complement(6181..6243)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=2  
/evidence=not\_experimental  
complement(6315..6389)  
/note="CDS is reported in Acc# AB017064  
gene\_id:MDN11.21  
unknown protein"  
/number=1  
/evidence=not\_experimental  
join(7906..8003,8102..8171,8284..9369,9442..10065)  
/note="contains similarity to bzp1 transcription factor  
gene\_id:MF21.2"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="BAA96993.1"  
/db\_xref="GI:8777403"  
/translation="METFLKDCSSVASSPISSPNISTLLKIKVLSMKETGLPASVH  
VRYCNKSPNLRKSLCAKSGYFKERDOLSEIIPQEPGARTPKIMLFTGCPVL  
IHPFNAGRCAGQFLMTRQSTGULCERFDIYNQVLIQNDITLVILKKQDLYP  
WSBDLIVSCIESLAFACMETLDEPKRERKPVIMLEGVNOPEYTIERLINDOT  
WIKDLTDLPPEPFCKIIGSLRQMKERYSPVALYASKSVIPEQNTDIIQRLAD  
LILTRDKAYRFVPGFYFACIALNLKHDYVLIQDQIVSLHTAOENFIYPAGNRQ  
VAPSOELTWESLFVYVSTESERHLTSSSNVRGKLMIDIFLSRLPYDEMKTTRPI  
ELIETVPMSPRESHDOLYLAVNAFLQVHNISQEEKSGTCSYVNCCKLQGEASLELVK  
NEKMPRLVYQALFIQQLNTHQAFKDCSCSFRFTNSADBSGSVYPSRPLTSQOSRCT  
DDETGPRNRPFLCFMQKDATLHDEFESTSPFIIHLEQDLVLSKSLSHSDNLKKNCG  
KRSASRNKTFGGQVTTACIGSVSFTSORKXANLQILRVNLFGRKTRSRXAGESE  
R"  
complement(join(10522..10773,10925..11443,11524..11826,  
11914..12027))  
/note="gene\_id:MF21.3"  
/codon\_start=1  
/evidence=not\_experimental  
/product="polygalacturonase"  
/protein\_id="BAA96994.1"  
/db\_xref="GI:8777404"  
/translation="MGRVHFGVSAPFVFCILGISANAKFTINISPPGSDITNALAKF  
NEACQFPITKSTVMI PKGYRKLGEIVVMGPKAFIRIALGFTVADGANGKEMVAFR  
NINFGKINGGVFDGEGNAMRVNNGKLPISIRIDFVTYPAKIRGITSLDK  
HFHINVIKAMVTFEDVKIIPAESPNTGDIHVRSDGIKINSFISTGDDCVSGDG  
MKMLVETVCGPGHGISIGLSGRYSHEENVSGIKIINCTLOETDNGLRDKTPVSAK  
TTTASDIHPENILILKVNSPILIDQVCECMNOCNKKPSTIKLANISPKIRGTSGNK  
DAVKLCSKGYPCQNVGVGDVNIQYTGADGPAIFQCSNVSPLKIVGQIRACSSPYTK  
PK"  
complement(join(14112..14562,14647..15668))  
/note="gene\_id:MF21.4"

Query Match 2.8%; Score 23; DB 8; Length 59372;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 592 GACGATTTGAGAGATGAGA 614  
 |||||  
 Db 38778 GACGATTTGAGAGATGAGA 38800

RESULT 7  
 AC116259  
 LOCUS  
 DEFINITION Rattus norvegicus clone CH230-15207, WORKING DRAFT SEQUENCE, 2  
 unorderd pieces.  
 ACCESSION AC116259  
 VERSION AC116259.8 GI:23194555  
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 1 (bases 1 to 232747)  
 Muzny,D,Marle., Metzker,M,lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alabrooks,S., Amin,A., Arguiano,D., Anyalbechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Bewald,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Detamo,C., Ding,Y., Dinh,H., DiVya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falle,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garra,W., Georgeogis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevar,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Huiyk,S., Hume,J., Idlebit,D., Jackson,A., Hollins,B., Howells,S., Hladun,S.L., Hodgson,A., Hogues,M., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensheva,L., Louised,H., Lozardo,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Mundaya,M., Murphy,M., Nair,L., Narkervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwokilemeh,O., Okwom,G., Olampunagoon,A., Pal,S., Parke,K., Patelernak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindecker,A., Popovic,D., Primus,E., Pu,L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Sherty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smajic,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Wilson,R., Wlezyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,D., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 232747)  
 AUTHORS Morley,K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 232747)  
 REFERENCE Rat Genome Sequencing Consortium.  
 AUTHORS Direct Submission  
 TITLE Submitted (11-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On Sep 19, 2002 this sequence version replaced gi:21745861.  
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GTNT  
 Center clone name: CH230-15207  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 22112 bases at least Q40  
 Consensus quality: 224872 bases at least Q30  
 Consensus quality: 225979 bases at least Q20  
 Estimated insert size: 229787; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 38020: contig of 38020 bp in length  
 \* 38021 38120: gap of unknown length  
 \* 38121 232747: contig of 194627 bp in length.  
 Location/Qualifiers  
 1 232747  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-15207"  
 5347..6361  
 /note="clone boundary  
 clone\_end:Sp6  
 site:ECORI  
 end\_sequence:BH336083"  
 5572..6089  
 /note="clone boundary  
 clone\_end:T7  
 site:ECORI  
 end\_sequence:BH336082"  
 30867..33254  
 /note="wgs\_end\_extension

misc\_feature  
 misc\_feature  
 misc\_feature

TITLE

misc\_feature clone\_end:T7"  
34423..38020  
/note="wgs\_end\_extension  
clone\_end:T7"  
misc\_feature 38121..39480  
/note="wgs\_end\_extension  
clone\_end:T7"  
BASE COUNT 69025 a 42956 c 44454 g 70267 t 6045 others  
ORIGIN

Query Match 2.8%; Score 23; DB 2; Length 232747;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 767 AAGAGTACAAATGATGATT 789  
|||||  
Db 214003 AAGAGTACAAATGATGATT 214025

RESULT 8  
AF146709 865 bp mRNA linear PLN 01-MAY-2000  
LOCUS  
DEFINITION Petunia axillaris An2 truncated protein (an2) mRNA, an2-S7 allele,  
complete cds.  
ACCESSION AF146709  
VERSION AF146709.1 GI:7673097  
KEYWORDS  
SOURCE Petunia axillaris  
ORGANISM Petunia axillaris  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; Lamiales; Solanales; Solanaceae; Petunia.  
REFERENCE 1 (bases 1 to 865)  
Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N.,  
Mol, J., and Koes, R.  
Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color  
Plant Cell 11 (8), 1433-1444 (1999)  
JOURNAL MEDLINE 99380006  
PUBMED 10449578  
REFERENCE 2 (bases 1 to 865)  
Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N.,  
Mol, J., and Koes, R.  
Direct Submission  
Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan  
1087, Amsterdam 1081HV, Netherlands  
LOCATION/Qualifiers  
FEATURES  
source 1..865  
/organism="Petunia axillaris"  
/mol\_type="mRNA"  
/cultivar="S7"  
/db\_xref="taxon:33119"  
/chromosome="6"  
/tissue\_type="petal limb"  
1..865  
/gene="an2"  
/allele="an2-S7"  
12..374  
/gene="an2"  
/note="anthocyanin regulator; myb-related transcription  
factor"  
/codon\_start=1  
/product="an2 truncated protein"  
/protein\_id="AA66734.1"  
/db\_xref="GI:7673098"  
/translation="MSTSNASTSGVRKGAATEEDLLIRECIKYGKMHLPVRAG  
LNRCKSCRLRLNLYLRPHIKRGDFSLDEVLLIRLHKLGNRWLSIAGRLPRTAND  
VKRWYNTDLRKKLAPHD"  
372  
variation  
/gene="an2"  
/note="nonsense mutation"  
/replace="c"  
BASE COUNT 305 a 136 c 195 g 229 t

ORIGIN

Query Match 2.7%; Score 22; DB 8; Length 865;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 TGAATTATCTAAGGCCACATAT 191  
|||||  
Db 181 TGAATTATCTAAGGCCACATAT 202

RESULT 9  
AF146707 868 bp mRNA linear PLN 01-MAY-2000  
LOCUS  
DEFINITION Petunia x hybrida An2 truncated protein (an2) mRNA, an2-W44 allele,  
complete cds.  
ACCESSION AF146707  
VERSION AF146707.1 GI:7673093  
KEYWORDS  
SOURCE Petunia x hybrida  
ORGANISM Petunia x hybrida  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; Lamiales; Solanales; Solanaceae; Petunia.  
REFERENCE 1 (bases 1 to 868)  
Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N.,  
Mol, J., and Koes, R.  
Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color  
Plant Cell 11 (8), 1433-1444 (1999)  
JOURNAL MEDLINE 99380006  
PUBMED 10449578  
REFERENCE 2 (bases 1 to 868)  
Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N.,  
Mol, J., and Koes, R.  
Direct Submission  
Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan  
1087, Amsterdam 1081HV, Netherlands  
LOCATION/Qualifiers  
FEATURES  
source 1..868  
/organism="Petunia x hybrida"  
/mol\_type="mRNA"  
/cultivar="W44"  
/db\_xref="taxon:4102"  
/chromosome="6"  
/tissue\_type="petal limb"  
1..868  
/gene="an2"  
/allele="an2-W44"  
12..395  
/gene="an2"  
/note="anthocyanin regulator; myb-related transcription  
factor"  
/codon\_start=1  
/product="an2 truncated protein"  
/protein\_id="AA66732.1"  
/db\_xref="GI:7673094"  
/translation="MSTSNASTSGVRKGAATEEDLLIRECIKYGKMHLPVRAG  
LNRCKSCRLRLNLYLRPHIKRGDFSLDEVLLIRLHKLGNRWLSIAGRLPRTAND  
VKRWYNTDLRKKLAPHDQKESKS"  
392..395  
variation  
/gene="an2"  
/note="causes frameshift leading to truncated protein"  
/replace="nnnn"  
BASE COUNT 307 a 137 c 195 g 229 t  
ORIGIN

Query Match 2.7%; Score 22; DB 8; Length 868;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 TGAATTATCTAAGGCCACATAT 191  
|||||

Db 181 TGAATTATCTAAGCCACATAT 202

RESULT 10  
AF146705 909 bp mRNA linear PLN 01-MAY-2000

LOCUS AF146705  
DEFINITION Petunia x hybrida An2 truncated protein (an2) mRNA, an2-W115  
allele, complete cds.

ACCESSION AF146705  
VERSION AF146705.1 GI:7673089

KEYWORDS  
ORGANISM Petunia x hybrida  
Petunia x hybrida  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; Lamiales; Solanales; Solanaceae; Petunia.

REFERENCE  
AUTHORS 1 (bases 1 to 909)  
Quattrocchio, F., Wang, J., van der Woude, K., Souer, E., de Vetten, N.,  
Mol, J., and Koes, R.  
TITLE Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color  
JOURNAL Plant Cell 11 (8), 1433-1444 (1999)

MEDLINE 99380006  
PUBMED 10449578  
2 (bases 1 to 909)  
Quattrocchio, F., Wang, J., van der Woude, K., Souer, E., de Vetten, N.,  
Mol, J., and Koes, R.  
REFERENCE  
AUTHORS Direct Submission  
Submitted (29-APR-1999) Genetics, Vrije Universiteit, De Boelelaan  
1087, Amsterdam 1081HV, Netherlands  
JOURNAL Location/Qualifiers

FEATURES  
source 1..909  
/organism="Petunia x hybrida"  
/mol\_type="mRNA"  
/cultiivar="W115"  
/db\_xref="taxon:4102"  
/chromosome="6"  
/tissue\_type="petal limb"  
1..909  
/gene="an2"  
/allele="an2-W115"  
12..401  
/gene="an2"  
/note="anthocyanin regulator; myb-related transcription  
factor"  
/codon\_start=1  
/product="An2 truncated protein"  
/protein\_id="AAF66730.1"  
/db\_xref="GI:7673090"  
/translation="MSTSNASTSGVRKGAWTEEDLLRECEIKYGEKMHLPVRAG  
LNRCKSCRLBMLVIRPHILKRGDPSLDEVELLIRLKLGLNRSLLAGRLPRTAND  
VKRYMTHLRKLLVPHDKOKSKTKP"  
391  
variation  
/gene="an2"  
/note="causes frameshift leading to truncated protein"  
/replace=""

BASE COUNT 324 a 137 c 199 g 248 t 1 others

ORIGIN

Query Match 2.7%; Score 22; DB 8; Length 909;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 170 TGAATTATCTAAGCCACATAT 191  
|||||  
181 TGAATTATCTAAGCCACATAT 202  
|||||

RESULT 11  
AC097536 194553 bp DNA linear PRI 12-JUN-2002

LOCUS AC097536  
DEFINITION Homo sapiens BAC clone RP11-803N5 from 2, complete sequence.  
AC097536 AC032034

VERSION AC097536.3 GI:18642970  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 194553)  
Sulston, J.E. and Waterston, R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)

MEDLINE 99063792  
PUBMED 9847074  
2 (bases 1 to 194553)  
Cadrioni, M., Haglund, K., and Creason, K.  
TITLE The sequence of Homo sapiens BAC clone RP11-803N5  
JOURNAL Unpublished (2001)  
REFERENCE  
AUTHORS 3 (bases 1 to 194553)  
Waterston, R.H.  
TITLE Direct Submission  
Submitted (18-OCT-2001) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
REFERENCE  
AUTHORS 4 (bases 1 to 194553)  
Waterston, R.H.  
TITLE Direct Submission  
Submitted (09-FEB-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
5 (bases 1 to 194553)  
Waterston, R.  
TITLE Direct Submission  
Submitted (12-JUN-2002) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Feb 9, 2002 this sequence version replaced gi:16502008.

COMMENT  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc  
Contact: saplens@wustl.wustl.edu  
----- Summary Statistics  
Center project name: H\_NH0803N05  
Drafting Center: WIBR  
-----

NOTICE: This sequence may not represent the entire insert of this  
clone. It may be shorter because we only sequence overlapping  
clone sections once, or longer because we provide a small overlap  
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. John D.  
McPherson, Department of Genetics, Washington University, St. Louis  
MO. For additional information about the map position of this  
sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
The RPC1-11 human BAC library was made from the blood of one male  
donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E.,  
Tateno, M., Catanesi, J.J., and de Jong, P.J. (1998) An improved  
approach for construction of bacterial artificial chromosome  
libraries. Genomics 51:1-8. The clone may be obtained either from  
Research Genetics, Inc. (<http://www.reegen.com>) or Pieter de Jong  
and coworkers at <http://www.chori.org>  
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP11-24K2, 2000 bp overlap; the clone sequenced to the right is RP11-8M15, 2000 bp overlap. Actual end is at base position 194553 of RP11-803M5.

Data from AC093400 was used to finish this clone, AC097536.

The sequence of AC032034 has been incorporated into AC097536.

#### FEATURES

source  
1. 194553  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="2"  
/clone="RP11-803M5"  
/clone\_1kb="RPCT-11"  
1026..1161  
/rpt\_family="MIR"  
2109..2183  
/rpt\_family="MIR"  
2859..2922  
/rpt\_family="L2"  
3889..3991  
/rpt\_family="L2"  
4065..4179  
/rpt\_family="ACHobo"  
5128..5310  
/rpt\_family="Alu"  
5469..5798  
/rpt\_family="L1"  
6813..8100  
/rpt\_family="ERV1"  
8383..8668  
/rpt\_family="Alu"  
8701..9167  
/rpt\_family="MALR"  
9273..9371  
/rpt\_family="L2"  
9372..9476  
/rpt\_family="G-rich"  
10860..10912  
/rpt\_family="AT-rich"  
11286..11310  
/rpt\_family="AT-rich"  
13348..13456  
/rpt\_family="L1"  
13457..14090  
/rpt\_family="L1"  
14305..15191  
/rpt\_family="L1"  
18364..18587  
/rpt\_family="L1"  
18630..19158  
/rpt\_family="MER1\_type"  
19171..19240  
/rpt\_family="L1"  
19241..19344  
/rpt\_family="Mariner"  
19345..19591  
/rpt\_family="L1"  
19607..19979  
/rpt\_family="MER1\_type"  
20085..20837  
/rpt\_family="L1"  
21033..21568  
/rpt\_family="L1"  
21771..22008  
/rpt\_family="MALR"  
22007..22147  
/rpt\_family="MALR"  
22241..22262  
/rpt\_family="AT-rich"

repeat\_region 22630..22686  
/rpt\_family="(TAGA)n"  
repeat\_region 22762..22813  
/rpt\_family="L1"  
repeat\_region 22814..23100  
/rpt\_family="Alu"  
repeat\_region 23101..23299  
/rpt\_family="L1"  
repeat\_region 25210..25357  
/rpt\_family="MIR"  
repeat\_region 25482..25638  
/rpt\_family="MIR"  
repeat\_region 25706..25815  
/rpt\_family="MIR"  
repeat\_region 25831..25851  
/rpt\_family="AT-rich"  
repeat\_region 26283..26357  
/rpt\_family="MALR"  
repeat\_region 26456..26721  
/rpt\_family="L1"  
repeat\_region 26767..27072  
/rpt\_family="Alu"  
repeat\_region 27512..27636  
/rpt\_family="MIR"  
repeat\_region 27870..27990  
/rpt\_family="ERV1"  
repeat\_region 28392..28436  
/rpt\_family="(CA)n"  
repeat\_region 28549..28676  
/rpt\_family="MER1\_type"  
repeat\_region 28677..28965  
/rpt\_family="Alu"  
repeat\_region 28966..29115  
/rpt\_family="MER1\_type"  
repeat\_region 29338..29878  
/rpt\_family="L1"  
repeat\_region 33682..34194  
/rpt\_family="L1"  
repeat\_region 34200..34630  
/rpt\_family="L1"  
repeat\_region 34631..34663  
/rpt\_family="(CA)n"  
repeat\_region 34664..34748  
/rpt\_family="L1"  
repeat\_region 37587..37899  
/rpt\_family="Alu"  
repeat\_region 39026..39148  
/rpt\_family="MALR"  
repeat\_region 39149..39461  
/rpt\_family="Alu"  
repeat\_region 39462..39722

Query Match 2.7%; Score 22; DB 9; Length 194553;  
Best Local Similarity 100.0%; Pred. No. 4.1;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 96 TGAAGGAAATGCACTTGT 117  
Db 19420 TGAAGGAAATGCACTTGT 19441

#### RESULT 12

LOCUS AC130926 219481 bp DNA linear HTG 15-NOV-2002  
DEFINITION Rattus norvegicus clone CH230-39011, WORKING DRAFT SEQUENCE.  
ACCESSION AC130926  
VERSION AC130926.3 GI:25007446  
KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 219481)

Murphy, D., Marie, M., Merz, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anyalbechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bialow, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., d'Souza, L., Davis, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Gant, R., Garcia, A., Garner, T., Garza, M., George, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guerra, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Joliver, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowls, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Louise, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahdarte, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinley, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokilemeh, O., Okwum, G., Olarnpusagoon, A., Pal, S., Parke, K., Paeternak, S., Paul, H., Perez, A., Perez, J., Pfannkuch, C., Plopper, F., Poldexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, B., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, M., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Soralle, R., Sosa, R., Steimle, M., Strong, R., Sutton, A., Swatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Vallas, R., Vera, V., Villasa, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, H., Wiley, K., Williams, G., Wilson, R., Wleciyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE  
JOURNAL

## REFERENCE

## AUTHORS

TITLE  
JOURNAL

## REFERENCE

## AUTHORS

TITLE  
JOURNAL

## COMMENT

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 15, 2002 this sequence version replaced gi:23264859.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

## ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

## ----- Project Information

Center project name: KBIN

Center clone name: CH230-39011

## ----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 213455 bases at least Q40

Consensus quality: 214651 bases at least Q30

Consensus quality: 215564 bases at least Q20

Estimated insert size: 224277; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 1 contigs. Gaps between the contigs

\* are represented as runs of N. The order of the pieces

\* is believed to be correct as given, however the sizes

\* of the gaps between them are based on estimates that have

\* provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

1 219481: contig of 219481 bp in length.

## Location/Qualifiers

1. 219481

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-39011"

698..25326

/note="clone\_boundary"

clone\_end:T7

site:

end\_sequence:BH286698"

complement(218245..218630)

/note="clone boundary"

clone\_end:Sp5

site:

end\_sequence:BH286728"

70629 a 42948 c 42163 g 60766 t 2975 others

## BASE COUNT

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 4.1;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 126574 AGAGATGAGAGCTTGAATT 126595

603 AGAGATGAGAGCTTGAATT 624

AC103077

AC103077

AC103077

AC103077

AC103077

AC103077

AC103077

AC103077

AC103077

AC103077

AC103077

AC103077

AC103077



REFERENCE 1 (bases 1 to 220494)  
 AUTHORS Mazy,D,Marle, Metzker,M, Lee, Abramzon,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D, Ayalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Biewald,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Butrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Diya,K, Diaper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,M, Guerra,W, Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M, Hollins,B, Howell,S, Hulik,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovat,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorenschweta,L, Louleseg,H, Lozado,R,J, Lu,X, Ma,J, Mesheswari,M, Mahindartine,M, Mahmoud,M, Malloy,K, Mangum,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mawhney,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Mllosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Muidasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nockelemehe,O, Okwomou,G, Olarnunagon,A, Pal,S, Parks,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C, Plopper,F, Polidexter,A, Popovic,D, Primus,E, Pu,L, L, Piazto,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rivers,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savary,G, Scherer,S, Scott,G, Shatman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajs,D, Sheed,A, Sodergren,E, Song,X,Z, Sorelle,R, Sosa,J, Steimle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K, Valas,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willison,R, Wleczky,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhausen,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R,A.

TITLE Unpublished  
 JOURNAL 2 (bases 1 to 220494)  
 REFERENCE Morley,K,C.  
 AUTHORS Direct Submission  
 TITLE Submitted (24-NOV-2001) Human Genome Sequencing Center, Department  
 JOURNAL of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 220494)  
 REFERENCE Rat Genome Sequencing Consortium.  
 AUTHORS Direct Submission  
 TITLE Submitted (23-NOV-2002) Human Genome Sequencing Center, Department  
 JOURNAL of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 COMMENT On Nov 23, 2002 this sequence version replaced gi:22855952.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GUDY  
 Center clone name: CH230-224J5  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 215732 bases at least Q40  
 Consensus quality: 217591 bases at least Q30  
 Consensus quality: 218727 bases at least Q20  
 Estimated insert size: 231920; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 6164: contig of 6164 bp in length  
 \* 6165 6264: gap of unknown length  
 \* 6265 216851: contig of 210587 bp in length  
 \* 216852 216951: gap of unknown length  
 \* 216952 218096: contig of 1145 bp in length  
 \* 218097 218196: gap of unknown length  
 \* 218197 219224: contig of 1028 bp in length  
 \* 219225 219324: gap of unknown length  
 \* 219325 220494: contig of 1170 bp in length.

## FEATURES

Location/Qualifiers  
 source  
 1..220494  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-224J5"  
 misc\_feature  
 1..3020  
 /note="wgs\_contig"  
 misc\_feature  
 6265..8020  
 /note="wgs\_contig"  
 BASE COUNT 74689 a 43966 c 42534 g 58135 t 1170 others  
 ORIGIN

Query Match 2.7%; Score 22; DB 2; Length 220494;  
 Best Local Similarity 100.0%; Pred. No. 4.1;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 603 AGAAGATGAAGAGCTTATTT 624  
 Db 70461 AGAAGATGAAGAGCTTATTT 70482

## RESULT 14

U32791 10913 bp DNA linear BCT 29-MAY-1998  
 LOCUS U32791/c  
 DEFINITION Haemophilus influenzae Rd section 106 of 163 of the complete  
 genome.  
 ACCESSION U32791 L42023  
 VERSION U32791.1 GI:1574662

## KEYWORDS

Haemophilus influenzae Rd  
 Haemophilus influenzae Rd  
 Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 Pasteurellaceae; Haemophilus.

## SOURCE

REFERENCE  
AUTHORS  
1 (bases 1 to 10913)  
Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A.,  
Kirkness,E.F., Kerlavage,A.R., Bilt,C.O., Tomb,J., Dougherty,B.A.,  
Merriam,J.P., McKenney,K., Sutton,G., Fitzhugh,W., Field,C.A.,  
Gocayne,J.D., Scott,J.D., Shiley,R., Liu,L.I., Glodek,A.,  
Cocayne,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E.,  
Cotton,M.D., Uterback,T., Hanna,M.C., Nguyen,D.T., Saudek,D.M.,  
Brandon,R.C., Fine,L.D., Fritchman,J.L., Fritchman,J.L.,  
Georghen,N.S., Gnehm,C.L., McDonald,L.A., Small,K.V., Frazer,C.M.,  
Smith,H.O., and Venter,J.C.  
Whole-genome random sequencing and assembly of Haemophilus  
influenzae Rd  
Science 269 (5223), 496-512 (1995)

JOURNAL  
MEDLINE  
7542800  
2 (bases 1 to 10913)  
Tatusov,R.L., Mushegian,A.R., Bork,P., Brown,N.P., Hayes,W.S.,  
Borodovsky,M., Rudd,K.E. and Koonin,E.V.  
Metabolism and evolution of Haemophilus influenzae deduced from a  
whole-genome comparison with Escherichia coli  
Curr. Biol. 6 (3), 279-291 (1996)

JOURNAL  
MEDLINE  
86398784  
3 (bases 1 to 10913)  
White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.  
Direct Submission  
Submitted (25-JUL-1995) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA

REFERENCE  
AUTHORS  
4 (bases 1 to 10913)  
White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.  
Direct Submission  
Submitted (27-SEP-1997) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
The H. influenzae sequence has been updated by R. Fleischmann. New  
database matches have been assigned, product names have been  
improved, and a number of frame shifts have been corrected. We  
gratefully acknowledge the work of Tatusov et. al. We have  
incorporated their annotation into the /notes fields of the  
corresponding H. influenzae genes

REFERENCE  
AUTHORS  
5 (bases 1 to 10913)  
White,O., Clayton,R.A., Kerlavage,A.R., Fleischmann,R.D.,  
Peterson,J., Hickey,E., Dodson,R. and Gwinn,M.  
Direct Submission  
Submitted (28-MAY-1998) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
The whole genome was shifted by 588 nucleotides for a new start  
On Oct 1, 1996 this sequence version replaced g1:1221866.  
Location/Qualifiers  
1..10913  
/organism="Haemophilus influenzae Rd"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:71421"  
complement(168..1295)  
/gene="HI1109"  
complement(168..1295)  
/note="similar to SP:P37389 PID:466706 GB:U00096  
PID:1789992 percent identity: 69.81; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="D-xylose ABC transporter, permease protein  
(xy1H)"  
/protein\_id="AAC22763.1"  
/db\_xref="GI:1574663"  
/translation="MFKLKSVNLQVYIMLAIAMVAFSVATDQAVLSARNISNLR  
QTSITGKLAIGMVFVISAIDLSVGLMGALADVMGFPLPVITATIALG  
LIFGIMGMWAVYKRVPSFIVTLAGYAFGLIGLGTGTVSPISGMVYIGGVYS  
DIAGYILGIAVIGFVLMGNQVRSROQLQVSAISKDPFKYALPAVYVIGALYILN  
DYRGIPFVLAVLAIALIGLFSKRTSGRHHVYIAGVITDAKISGINVETKALITRA  
MNGVIALAGLILSARLGAQSPSAGQNAELDAIACVIGASLAGVSGVGVVIGAL  
IIASLDNGMSMLDPTFWQYIVKGIILLAIVMTIDTSNKKK"  
complement(1299..2810)  
gene

CDS  
/gene="HI1110"  
complement(1299..2810)  
/note="similar to SP:P37389 PID:466705 GB:U00096  
PID:1789991 percent identity: 71.46; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="D-xylose ABC transporter, ATP-binding protein  
(xy1G)"  
/protein\_id="AAC22764.1"  
/db\_xref="GI:1574664"  
/translation="MALLEMKHTKKFGDVTALNISIEHAGEIISLCEGSGSKT  
IMKILGIVCPGDIYFSESELKARNIRDTREKGISIHQELTLVKMSYLENF  
LGNETHIKGILADNEMVLRKMLQVOLDADNDPTVRGELGQOOLVETAKLVNKF  
RLILDEPRLSLTEKETELINLIKOLKANINACIYISHKTNVKSADKICVIRGE  
HVTQKASITVEDDITTMVGREITSLYEPHEIKDEILVENLSAMPHINIKYV  
DNYSFSLHEEILGIVAGLVSGRTDVOVCLFSGYEGKEGNIIFINQKQVINKCAQAI  
EKKIIVWPEDEKRGIVISWVGKNITLSLKSICYCGKMGVNAKEKQVIGSAIKRLK  
VKTSEPDLPGRLSGNOQKAILAKSLPKLILDEPRGIDVGAKYIYKLINQL  
AQEGIALIVISSELPEVIGISDRVLVWQKCLKASLINTALQEQWETALKE"  
complement(2861..3859)  
/gene="HI1111"  
complement(2861..3859)  
/note="similar to SP:P37387 PID:466704 GB:U00096  
PID:1789990 percent identity: 75.99; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="D-xylose ABC transporter, periplasmic-binding  
protein (xy1F)"  
/protein\_id="AAC22765.1"  
/db\_xref="GI:1574665"  
/translation="MKIKSLILTVGALTYFSSASAKDKIGLSIDDLRLERQKDR  
KDIVNASEMGAVFVQASANGDSAQISQENINKNIDVLTVPNHEGLVSNVISA  
KKEGKIVLADRLINNADLPYVSFNEKVGELQAKSIVAKPEGNTFLGSGSPVDN  
AKIFRKGOMVLDPLIASGKIKVYGQVDSWLAERALKQIMENALTANKNNVAVAS  
NDATAGGAIQALSAQGLSGKVAISGQADALAKRIYNGSQTWTVVYPIRLKDKAE  
IATVGLKNEKTEANAEINLNLKNVPAYLDPIAVDRKNINETYIKQGFHKESLYH"  
4093..5412  
/gene="HI1112"  
4093..5412  
/note="similar to GB:X04691 SP:P00944 GB:X01996 GB:S68257  
PID:148279 percent identity: 71.30; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="xylose isomerase (xy1A)"  
/protein\_id="AAC22766.1"  
/db\_xref="GI:1574666"  
/translation="MTTFPKIKISREKGSNTPAFKHYDANOVILKTMAEHLRL  
AVCYWHTFCWNGNDMFGLSLERSWQNSNMLAGAEKADIAEFKLKLVPIYCFHD  
VDIAPBEGNSVREYVONFHHIVDILERQVETGVLMTGNTNCFNTPRYMSGAATNPB  
EVPAMATQVFNAMNATORLGENYVLMGREGYETLLNDLREBQIRFQVQVVE  
HKKIKGKGLILPEKQDEPTKHOYDVAVTVGKQFQGLKEIKVNIAMNATHTG  
HTQHELATKACDIDIGSIDANRGDQDLPQFNSVENTLWVEILIKHGGFTTG  
GFNPDAKIRQSIDPDIYFAHIGALIDVALSLKRAKMLQETLQKIVARVAGMS  
ELGQHLIQGKTSLETLAQVQKDLAKPVSQGOEYLENLVNDVITS"  
5415..6950  
/gene="HI1113"  
5415..6950  
/note="similar to PID:2293419 percent identity: 52.56;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="xyulose kinase"  
/protein\_id="AAC22767.1"  
/db\_xref="GI:1574667"  
/translation="MVKQCGCIPHPFILDIVYIGIDCQGTQKALVDSVQKVIK  
GVYAKHELTITQSNGRBQDQPNWMIIEALQALQIALKQAKNSPHSPMLVKGIGISGOO  
complement(1299..2810)  
gene

HG1VMDXNDRLPYKAKLWCDTETATENDILIEKLGSGQAVPEKLGITICOTGYTASKL  
 SWFRQNPDKFANI RIKIMLPHDYLANWLTGKCTEFEDASGSGFVVKREKREK  
 YLAPBELMDEPDLKLSABOKIYGLKEITLGFENENIVSTGGNNMGAIGTGI  
 REGIATWSLGTGLYATOKPLNPLNPMANCSNNGLPIVAVNNITSSKOLAN  
 LNIIDEENLOAQAPIGANGITLIPFGGEVVPYLPNKASIIIGDSENPRNIC  
 RAMESSTFLIRYGLDIFROAGLKTSGIRIIGGAKSSFRQMIADVMSEVCLQOE  
 EAAALGGAIOAMWANGEGELEFCEFTIHDENSKAVPNLSQVKNYQNAVERYLTHLS  
 QLYA  
 gene complement (7001. .7927)  
 /gene="H1114"  
 complement (7001. .7927)  
 /gene="H1114"  
 /note="similar to SP:P17963 GB:M33577 GB:X54492 PID:147576  
 PID:41763 percent identity: 78.25; identified by sequence  
 similarity: putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="ADP-L-glycero-D-mannohexose-6-epimerase (f6ad)"  
 /protein\_id="AAC22768.1"  
 /db\_xref="GI:1574668"  
 /translation="MIIVTGAGFISNIYKALNDGRKDIIVVNLKDGTRFANLVD  
 LDYADVCKEDFIASIGDERGIDAVFHGCSATBEDCKYIMHNNVEYKELH  
 YCIDRELPFYASATYGTQVFREREBGLNVTGYSKFLPDQYRNILPEASP  
 VCGFRFNVYGPENHKGMSAVAFILNOLKGENPKLFGSGHFERDFVYGVDAV  
 VNIWQNGISGIYNLGTGNAESFRAVADVVFHKGGEIETIIPFPHLSRYOEYTO  
 ADLTKLSTGYDKPFTVAGVTGYMAMLRK"  
 complement (7980. .8483)  
 /gene="H1115"  
 complement (7980. .8483)  
 /gene="H1115"  
 /note="similar to PID:1213065 SP:P52229 percent identity:  
 36.36; identified by sequence similarity: putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="thioredoxin, putative"  
 /protein\_id="AAC22769.1"  
 /db\_xref="GI:1574669"  
 /translation="MKIKLLKNGLSLPLFEIYVITSLDFVRPPVPEEINKITLDL  
 QGNFSLSELDONKPTLLIFMGTCGCRYSPLAKENKQVVSVALRSGNEADY  
 NDYLSKNDHYFTVNDPKGFARMQINVTPTVILSKMDVTTGLTISYKGLKRL  
 FFAFPG"  
 gene 8594. .9265  
 /gene="H1116"  
 CDS 8594. .9265  
 Query Match 2.5%; Score 21; DB 1; Length 10913;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 788 TTCTGCTGAATGACTTAT 808  
 Db 1114 TTCTGCTGAATGACTTAT 1094  
 RESULT 15  
 AE001227/c 26849 bp DNA linear BCT 16-JUL-1998  
 DEFINITION Treponema pallidum section 43 of 87 of the complete genome.  
 ACCESSION AE001227 AE000520  
 VERSION AE001227.1 GI:3322797  
 KEYWORDS  
 SOURCE Treponema pallidum  
 ORGANISM Treponema pallidum  
 Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
 REFERENCE  
 1 (bases 1 to 26849)  
 Fraser, C.M., Norris, S.J., Weinstein, G.M., White, O., Sutton, G.G.,  
 Dodson, R., Swin, M., Hickey, E.K., Clayton, R., Ketchum, K.A.,  
 Sodergren, E., Hardham, J.M., McLeod, M.P., Salzberg, S., Peterson, J.,  
 Khalak, H., Richardson, D., Howell, J.K., Chidambaram, M.,  
 Uetzerback, T., McDonald, L., Artlich, P., Bowman, C., Cotton, M.D.,  
 Venter, J.C. et al.  
 Complete genome sequence of Treponema pallidum, the syphilis  
 spirochete

JOURNAL Science 281 (5375), 375-388 (1998)  
 MEDLINE 98332770  
 PUBMED 9665876  
 REFERENCE 2 (bases 1 to 26849)  
 AUTHORS Fraser, C.M., Norris, S.J., Weinstein, G.M., White, O., Sutton, G.G.,  
 Dodson, R., Swin, M., Hickey, E.K., Clayton, R., Ketchum, K.A.,  
 Sodergren, E., Hardham, J.M., McLeod, M.P., Salzberg, S., Peterson, J.,  
 Khalak, H., Richardson, D., Howell, J.K., Chidambaram, M.,  
 Uetzerback, T., McDonald, L., Artlich, P., Bowman, C., Cotton, M.D.,  
 Fujii, C., Garland, S., Hatch, B., Horst, K., Roberts, K., Matchey, L.,  
 Weidman, J., Smith, H.O. and Venter, J.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-MAR-1998) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 FEATURES  
 source  
 1. 26849  
 /organism="Treponema pallidum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:160"  
 211. .283  
 /gene="tRNA-Lys-1"  
 211. .283  
 /gene="tRNA-Lys-1"  
 /product="tRNA-Lys"  
 /note="codon recognized: AAG"  
 343. .423  
 /gene="tRNA-Lys-4"  
 343. .423  
 /gene="tRNA-Lys-4"  
 /product="tRNA-Lys"  
 /note="codon recognized: CUA"  
 434. .505  
 /gene="tRNA-Gly-2"  
 434. .505  
 /gene="tRNA-Gly-2"  
 /product="tRNA-Gly"  
 /note="codon recognized: GGC"  
 586. .1914  
 /gene="TP0506"  
 586. .1914  
 /gene="TP0506"  
 /note="similar to GB:AE000783 percent identity: 25.17;  
 identified by sequence similarity: putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="trigger factor (tig)"  
 /protein\_id="AAC65494.1"  
 /db\_xref="GI:3322798"  
 /translation="MELQKFTALAQOYELVAVARDAORHYORFVEYERRARLP  
 GRKGVPLAVLERKYGSAIRODAALAMEKLEEGFAQASODSOLPISRPILKXPP  
 VPPDDPSFPAVIVDFPSEVELANTSGFSLSPPTVSYTEEDVSRELTRIOERNALYTD  
 KADSCAEVGDIAIVDYHEVDSGAVRPGTERGAVFTYGVESGFPALGODLIGMKLG  
 ORCLFAKRAQMLKDEAQAQVRLKALKQRLPBLDDELADQVSDAFRTLDLITRSVRO  
 NLALEALAEHXYRQLRLIVRENPSLPEBLVGVEMESRWALVMROFGVSLGTP  
 ONKLOFQOORPVEVEELKORVIVELIKOEOVSVAEELFEYVRIASATGSKERV  
 REYVAGEERBRALCEGIRREKLCQKILGRCYTECGESGLTDELQDSNA"  
 1907. .2518  
 /gene="TP0507"  
 1907. .2518  
 /gene="TP0507"  
 /note="similar to GB:L42023 SP:P43867 PID:11005637  
 PID:1220799 PID:1204962 percent identity: 60.42;  
 identified by sequence similarity: putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="ATP-dependent Clp protease proteolytic component  
 (clp-1)"  
 /protein\_id="AAC65495.1"  
 /db\_xref="GI:3322799"  
 /translation="MEERENLVPIYVTEGGGERSYDIFSRLLKRRITFVDEITDA  
 VADLVVAQLFLESQNPDKDISLYNSPGAAVAGVAGVATQHCPEVQITCLGQAS  
 SPAAVILAGAGARKFALPSSRMVHIQPGVQGVQGSADVCIQOEILRLKTLITIAVFA  
 LHTGSEQVREDMERDFTLSAQACSYSIVDTVMKRRKAGV"

gene 2505..3752  
/gene="TP0508"  
CDS 2505..3752  
/gene="TP0508"  
/note="similar to GB:L18867 SP:P3138 GB:Z23278 PID:347729  
PID:389977 percent identity: 56.78; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="ATP-dependent Clp protease subunit X (clpX)"  
/protein\_id="AAC65496.1"  
/db\_xref="GI:3322800"  
/translation="MLRSKGDVLVGCSCGCKEDERRRIYTGHSICNYCVERCAEY  
LBRKPSALAMTKEIPTPELKAADQVYIGODLAKRLVAVVNHRYAGRSID  
IDVLEKSNVLLIGPTSGSKTLAKTSOMKVPALADATLTTEGVGEVENT  
LKUVONANGVALAERGIITIDEIDKSRSENVSTIRDSGSGVOALAKIIEGTA  
SVPPQGRKHNDQMLRVDTSNLIETCGAPVGDGIVGTRCNRPVGFADVATVE  
RGQLMHEVDVLPDLVKGFLPEIIGRLPVYALDALSKEDELNIVRPNALVROPE  
ALFALDQVRLVFEDDALDAIAQAIDQGTGARGISIVERLMDAMEASLKGKEL  
CITKKVYTOEEKASVRLVSERTA"  
3849..4415  
/gene="TP0509"  
CDS 3849..4415  
/gene="TP0509"  
/note="similar to PID:1064782 GB:AL009126 percent  
identity: 48.92; identified by sequence similarity;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="alkyl hydroperoxide reductase (ahpC)"  
/protein\_id="AAC65497.1"  
/db\_xref="GI:3322801"  
/translation="MESLIGKRVIDFLPAVVGSKTEVSNASIKSMAVPMFPADP  
TPVCPTELDARVYPSFVELGCKVYSVSDSEYVHAMADATDTITKLYEMISDA  
GKLAFPGVLLPDTWHLRGTFFVDPGLVAEFVHDWIGRDADELRRVQSAQFVA  
KHGQVCPARKPKGAKTLKPGIDLIGKI"  
complement(4596..6413)  
/gene="TP0510"  
complement(4596..6413)  
/gene="TP0510"  
/note="similar to GB:AE000783 percent identity: 59.57;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="GTP-binding membrane protein (lepa)"  
/protein\_id="AAC65498.1"  
/db\_xref="GI:3322802"  
/translation="MRMACVRRVNFCTVAHIDGKSTLADRLIERTRAVEBRLQA  
QMTNMLERRRGITIKSHAVCIPTDAHGEVLANPDPGADPAVEYRAIACB  
GALLVDAOTGVESQTSNLVLEHNLLEIIPVINKIDLPADVPRLQOVENHLDG  
PASVLSATGENVDAIFDAITRIIPROSGTAALQALVPOCHVDQGVVHTRIV  
FEGVITSGNVIKRENSNGAEYREVEITGVFENLAKREALCAGDVGISANVYTSDDV  
GDTITDASCPCDTPRAGFRVRYKVPVSSVVDVDECEQLREALERLANDASISMR  
DSIALGHCRCGFLGLHLLEVQOLREBNQVITPAQOVYVFLKGTGORIVCN  
PAHYPLEOEIAQVHEPIRATITPTVELGAVMTLCIEKRAYOTAANVLDQKVELY  
EMPIAELFGVYDLKSIHSGVASFDEVELIESKLTDLVKVDILNGKVPDALQCYR  
PHARRAQAVCARLKEISRQFKIAIQSGGQIISRETVSPRKVYLAKCYGDDIT  
RKRRLKQKQKGRMVGDEIPTQAFSLKEASDA"  
6456..7082  
/gene="TP0511"  
CDS 6456..7082  
/gene="TP0511"  
/note="similar to GP:1881322 percent identity: 31.79;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="transcription factor (card)"  
/protein\_id="AAC65499.1"  
/db\_xref="GI:3322803"

gene 7127..8326  
/gene="TP0512"  
CDS 7127..8326  
/gene="TP0512"  
/note="similar to GB:X72382 SP:Q08113 PID:313163 percent  
identity: 36.31; identified by sequence similarity;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAC65500.1"  
/db\_xref="GI:3322804"  
/translation="MRGACVQKKEYLPDTSRQPCVCLISLTVBALEARSFELVV  
TVPAGVAAESQVACDSRLSAFSPRTRPVILVPGAHRSASVRAGLDAMAATHADV  
VLVHDCARFVSVALHSHVLEATRCYGAAPVLEADTPKGVADGSIETHLIRSRVR  
LAQTPGFCYAS/CAAHRAATDGEQYTDSELVARYGGVHVCAGERSVKLTYPED  
LEBRSAPLRTREISVLPCTERBGLAVGCTDMHALCAGRPLIAGIHS PKKGAQGH  
SDADVLAHNSIDALGAAGLGDIGTFPSCDGRKMDASHCALKLRHTQVLRACRMLV  
MLDAVVCLEQPALHPREARMASLAQALDTHVTRVFKAKTARLGPVSGAAVTAVQ  
VVLAKKI"  
8400..9800  
/gene="TP0513"  
CDS 8400..9800  
/gene="TP0513"  
/note="similar to SP:P23868 GB:X52114 PID:43137 PID:443993  
PID:606224 percent identity: 23.86; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="K+ transport protein (trkA)"  
/protein\_id="AAC65501.1"  
/db\_xref="GI:3322805"  
/translation="MRVIVGCGCTNGRELTRRLSARGHQTTLVAAHEPVPRKI PAGA  
OALVROKHORATLGERAGIAKQAFAVAVTEVEDELNLVGLIADILNAGVYKIKOTPD  
HYAHNCRSRCTFGIDPLISADKEMRAVSVTEGALSDVPLKNAPEYELARPIA  
KSGGLDITLSDRRRLVKISFVAVAFVGRSVIPSGETMLASGRLSYLCAERHNGR  
FELAGFKIHPVKIKALIGMSAVGTLVADVAEKCKPFFSSAFSLSPERRASLVLD

Query Match 2.5%; Score 21; DB 1; Length 26849;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 424 CCTCAACGACGCAAGTATTC 444  
DB 5442 CCTCAACGACGCAAGTATTC 5422

Search completed: January 30, 2004, 10:14:24  
Job time : 2286 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 08:32:02 ; Search time 241 Seconds  
(without alignments)  
9240.815 Million cell updates/sec

Title: US-10-033-190-1

Perfect score: 825

Sequence: 1 atgacagatcatcatgtc.....tatgaatcattgtattaa 825

Scoring table: OLIGO\_NNC

Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : N\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	825	100.0	825	24	ABO73046 Tomato anthocyanin
2	567	68.7	1012	24	ABO73048 Tomato anthocyanin
3	23	2.8	1300	24	ABK65274 Arabidopsis cDNA e
4	21	2.5	300	21	AAA01109 Human colon cancer
5	21	2.5	390	25	ABX20395 Human GIP-mannose
6	21	2.5	765	21	AAA02459 Human colon cancer
7	21	2.5	1830121	17	AA742063 Haemophilus influenza
8	20	2.4	741	24	ABZ12316 Arabidopsis thaliana

9	20	2.4	859	21	AA041093 Arabidopsis thaliana
10	20	2.4	684707	24	ABO67196 Listeria innocua C
11	20	2.4	3011208	24	ABO69245 Listeria innocua D
12	19	2.3	464	22	AAK62944 Human immune/haema
13	19	2.3	737	22	AA522921 DNA encoding novel
14	19	2.3	810	21	AAA70171 Plasmodium falcipar
15	19	2.3	1137	21	AA050784 Arabidopsis thaliana
16	19	2.3	1390	21	AA050357 Arabidopsis thaliana
17	19	2.3	1392	21	AA050357 Arabidopsis thaliana
18	19	2.3	1549	22	AAK83785 Arabidopsis thaliana
19	19	2.3	1549	22	AAK83785 Arabidopsis thaliana
20	19	2.3	1911	20	AAK83786 Human immune/haema
21	19	2.3	2493	23	ABX178079 Human immune/haema
22	19	2.3	2722	21	AA085057 DNA comprising a b
23	19	2.3	2722	21	AA085057 DNA construct comp
24	19	2.3	7153	22	AAK67687 Human immune/haema
25	19	2.3	13469	22	AA046593 Tumour suppressor
26	18	2.2	41	24	ABO77712 IGF binding protein
27	18	2.2	41	24	ABN86193 Insulinoid growth
28	18	2.2	95	22	ABA50855 Human breast cell
29	18	2.2	95	22	ABA68828 Human foetal liver
30	18	2.2	95	22	ABA35780 Probe #14246 for g
31	18	2.2	95	22	AAK17164 Human brain expres
32	18	2.2	95	22	AAK42949 Human bone marrow
33	18	2.2	95	22	AA123713 Probe #13646 for g
34	18	2.2	95	22	AA149026 Probe #17712 used
35	18	2.2	95	22	AA109331 Probe #9322 used t
36	18	2.2	95	23	AB042581 Human liver bingie
37	18	2.2	95	24	AB017011 Human genome-deriv
38	18	2.2	306	24	ABN66897 Streptococcus poly
39	18	2.2	360	24	AA030471 Human NKp46D1 (iso
40	18	2.2	432	22	ABA45736 Human breast cell
41	18	2.2	432	22	ABA56245 Human foetal liver
42	18	2.2	432	22	ABA25884 Probe #4350 for ge
43	18	2.2	432	22	AA04428 Human brain expres
44	18	2.2	432	22	AAK29924 Human bone marrow
45	18	2.2	432	22	AA114514 Probe #4447 for ge
46	18	2.2	432	22	AA135889 Probe #4575 used t
47	18	2.2	432	22	AA104340 Probe #4331 used t
48	18	2.2	432	23	AB029573 Human liver single
49	18	2.2	432	24	AB044492 Human genome-deriv
50	18	2.2	432	24	AB193882 Arabidopsis thaliana
51	18	2.2	432	24	AB193882 Arabidopsis thaliana
52	18	2.2	547	25	ABX57220 Arabidopsis thaliana
53	18	2.2	612	24	AB066283 Arabidopsis thaliana
54	18	2.2	612	24	ABK65317 Arabidopsis cDNA e
55	18	2.2	711	24	AA030464 Human NKp46 (iso
56	18	2.2	750	21	AA042691 Arabidopsis thaliana
57	18	2.2	750	24	AB212526 Arabidopsis thaliana
58	18	2.2	762	24	AA030459 Human NKp46 (iso
59	18	2.2	788	24	ABO76567 C. albicans BAX-as
60	18	2.2	901	22	AA085189 Nucleotide sequenc
61	18	2.2	952	21	AA045824 Arabidopsis thaliana
62	18	2.2	955	21	AA034136 Arabidopsis thaliana
63	18	2.2	982	24	ABN98500 Arabidopsis thaliana
64	18	2.2	994	22	AA006658 A. thaliana transc
65	18	2.2	999	21	AA234942 Wheat anti-silenci
66	18	2.2	1033	22	AA005751 Arabidopsis thaliana
67	18	2.2	1033	24	ABK65155 Arabidopsis cDNA e
68	18	2.2	1033	24	ABK14222 A. thaliana Produc
69	18	2.2	1033	25	AA047497 Arabidopsis thaliana
70	18	2.2	1043	22	AA085191 Nucleotide sequenc
71	18	2.2	1083	23	AA073889 DNA encoding novel
72	18	2.2	1135	21	AA059159 Human secreted pro
73	18	2.2	1135	25	AB273315 Secreted protein-e
74	18	2.2	1135	25	AB266928 Human secreted pro
75	18	2.2	1141	24	AB070816 Myb-related trans
76	18	2.2	1161	22	AA005758 Arabidopsis thaliana
77	18	2.2	1266	21	AA081130 Human secreted pro
78	18	2.2	1266	21	ABO68021 Human secreted pro
79	18	2.2	1266	24	ABO69843 Listeria monocytog
80	18	2.2	1322	22	AA005782 Arabidopsis thaliana
81	18	2.2	1322	24	ABK65334 Arabidopsis cDNA e

82	18	2.2	1322	25	AA047525	Arabidopsis thalia	155	17	2.1	565	25	AB217841	S2 subtraction lib
83	18	2.2	1416	24	AA030465	Human NKp46 (Isofo	156	17	2.1	604	24	ABX65526	Helicobacter pylor
84	18	2.2	1467	24	AA030461	Human NKp46 (Isofo	157	17	2.1	606	24	ABX66835	Helicobacter pylor
85	18	2.2	1648	22	AA034687	Human DNA for a no	158	17	2.1	607	23	ABV58521	Human prostate exp
86	18	2.2	1648	22	AA034688	Human DNA for a no	159	17	2.1	615	24	ABN68451	Streptococcus poly
87	18	2.2	2175	24	AB077707	Insulin-like growt	160	17	2.1	620	19	AA013951	H. pylori GHPO 58
88	18	2.2	2175	24	ABN66188	Human Insulinoid g	161	17	2.1	620	22	AAK87194	Human immune/haema
89	18	2.2	2208	25	ABT14994	Pathogen specific	162	17	2.1	648	24	ABX66325	Helicobacter pylor
90	18	2.2	2256	22	AAH15044	Human cDNA sequenc	163	17	2.1	650	22	AAH50992	Human nGPCR36 codi
91	18	2.2	2337	24	ABN85768	Arabidopsis yellow	164	17	2.1	660	24	AB070225	DNA encoding human
92	18	2.2	2787	24	AB232289	Candida albicans e	165	17	2.1	667	24	ABX65734	Helicobacter pylor
93	18	2.2	3012	22	AAH98552	Human EST-derived	166	17	2.1	677	24	AA062889	Human G protein-co
94	18	2.2	3335	21	AA096708	Reporter gene PST1	167	17	2.1	693	25	AB252484	Aspergillus oryzae
95	18	2.2	3458	23	ABL10801	Drosophila melanog	168	17	2.1	715	22	AAK75358	Human immune/haema
96	18	2.2	4310	23	ABL25840	Drosophila melanog	169	17	2.1	715	22	AAK75359	Human immune/haema
97	18	2.2	4834	23	ABL12040	Drosophila melanog	170	17	2.1	715	22	AAK75360	Human immune/haema
98	18	2.2	4865	23	ABL02908	Drosophila melanog	171	17	2.1	731	23	ABV20846	Human prostate exp
99	18	2.2	5575	23	ABL29660	Drosophila melanog	172	17	2.1	731	23	ABV26693	Human prostate exp
100	18	2.2	6563	23	ABL17835	Drosophila melanog	173	17	2.1	746	24	AB052818	Oligonucleotide fo
101	18	2.2	8353	23	ABL15774	Drosophila melanog	174	17	2.1	746	24	AB052819	Oligonucleotide fo
102	18	2.2	9513	23	AA036311	Human cardiovascular	175	17	2.1	764	22	AAH0412	Nucleotide sequenc
103	18	2.2	10917	19	AAV25568	Bacillus subtilis	176	17	2.1	765	22	AAH07306	Human cDNA clone (
104	18	2.2	11427	22	AA036312	Human cardiovascular	177	17	2.1	776	24	ABK47742	CDNA encoding huma
105	18	2.2	15827	23	ABL09558	Drosophila melanog	178	17	2.1	791	24	ABN99077	Arabidopsis thalia
106	18	2.2	15951	24	ABL70374	Chemically treated	179	17	2.1	795	24	ABN99065	Arabidopsis thalia
107	18	2.2	15951	24	ABL33681	Human immune syste	180	17	2.1	804	21	AA045885	Human prostate exp
108	18	2.2	15951	24	ABL33681	Human immune syste	181	17	2.1	811	23	ABV11480	H. pylori cytoplasm
109	18	2.2	17674	24	ABL33345	Human immune syste	182	17	2.1	828	19	AAK30633	H. pylori cytoplasm
110	18	2.2	23203	23	ABL17834	Drosophila melanog	183	17	2.1	837	19	AAK30634	Human immune/haema
111	18	2.2	25577	22	AAK65015	Human immune/haema	184	17	2.1	848	22	AAK62731	Human immune/haema
112	18	2.2	25577	22	AAK79382	Human immune/haema	185	17	2.1	849	25	AB237832	N. gonorrhoeae nuc
113	18	2.2	32191	25	ABX74033	Human genomic DNA	186	17	2.1	849	25	AB239338	Helicobacter pylor
114	18	2.2	32191	25	ABX74033	Human novel polyinu	187	17	2.1	865	24	ABX66986	Arabidopsis thalia
115	18	2.2	49561	22	AAK82012	Human immune/haema	188	17	2.1	868	21	AA049989	Arabidopsis thalia
116	18	2.2	175737	22	AAK83571	Human cDNA diftere	189	17	2.1	871	21	AA034284	Arabidopsis thalia
117	18	2.2	168218	25	AB084281	Chromosome 2 Aschim	190	17	2.1	888	24	AB213785	Arabidopsis thalia
118	18	2.2	1082138	21	AA022305	Arabidopsis thalia	191	17	2.1	888	24	ABN98734	Arabidopsis thalia
119	18	2.2	2155561	24	ABN71557	Streptococcus poly	192	17	2.1	890	22	AAH03553	Human cDNA clone (
120	18	2.2	2155561	24	ABN90521	Genomic sequence o	193	17	2.1	895	21	AA033660	Arabidopsis thalia
121	18	2.2	2944528	24	ABN03041	Listeria monocytog	194	17	2.1	897	24	AB213179	Arabidopsis thalia
122	17	2.1	161	21	AA02500	Human secreted pro	195	17	2.1	897	24	AA018020	Human cDNA encodin
123	17	2.1	219	22	AAH77910	Nucleotide sequenc	196	17	2.1	933	24	AB214438	Arabidopsis thalia
124	17	2.1	232	22	AA051125	Human breast cell	197	17	2.1	933	24	AA098117	Human DNA for pote
125	17	2.1	232	22	AAK17429	Human brain expres	198	17	2.1	960	24	AB214333	Arabidopsis thalia
126	17	2.1	232	23	AB042861	Human liver single	199	17	2.1	963	24	AA098110	Human DNA for pote
127	17	2.1	236	24	AB098448	Human ORF255 codin	200	17	2.1	974	24	ABN98513	Arabidopsis thalia
128	17	2.1	271	25	ABX85125	Corn ear-derived p	201	17	2.1	978	22	AA026569	Human cDNA encodin
129	17	2.1	282	25	ABX30719	Human GDP-mannose	202	17	2.1	978	25	ABX73910	Human novel polyinu
130	17	2.1	300	21	AA00585	Human colon cancer	203	17	2.1	990	19	AAK14165	H. pylori GHPO 542
131	17	2.1	303	22	AAH77908	Nucleotide sequenc	204	17	2.1	994	24	AA018019	Human cDNA encodin
132	17	2.1	312	21	AA045319	Arabidopsis thalia	205	17	2.1	1059	24	AB213683	Arabidopsis thalia
133	17	2.1	315	24	ABN19961	Human ORFX polynuc	206	17	2.1	1065	21	AA070094	Plasmodium falcipa
134	17	2.1	328	21	AAA42607	Human secreted exp	207	17	2.1	1091	24	ABK35688	CDNA sequence #89
135	17	2.1	330	24	ABN77319	Human ORF2266 cDNA	208	17	2.1	1104	11	AA004523	Lipoxigenase gene.
136	17	2.1	363	21	AA018316	Human secreted pro	209	17	2.1	1120	21	AA046213	Arabidopsis thalia
137	17	2.1	365	25	AB217736	S2 subtraction lib	210	17	2.1	1123	21	AA035071	Arabidopsis thalia
138	17	2.1	397	25	ABX62754	Arabidopsis thalia	211	17	2.1	1125	24	AB214528	Nucleotide sequenc
139	17	2.1	413	25	ABX53237	Bovine EST associa	212	17	2.1	1139	22	AA080400	Arabidopsis thalia
140	17	2.1	419	24	ABL40994	Human neutregulin 2	213	17	2.1	1153	24	AA031884	Lactobacillus rham
141	17	2.1	420	23	AA076562	DNA encoding novel	214	17	2.1	1157	21	AA040127	Arabidopsis thalia
142	17	2.1	456	22	AB046009	Human breast cell	215	17	2.1	1160	21	AA038848	Arabidopsis thalia
143	17	2.1	456	22	AAK04695	Human brain expres	216	17	2.1	1168	22	AA199212	Human excretory re
144	17	2.1	456	23	AB029859	Human liver single	217	17	2.1	1168	22	AA199213	Human excretory re
145	17	2.1	470	24	ABX65996	Helicobacter pylor	218	17	2.1	1168	22	AA163562	Human kidney relat
146	17	2.1	482	21	AA038143	Zeal maye DNA fragm	219	17	2.1	1168	22	AA163563	Human kidney relat
147	17	2.1	510	24	ABX66567	Helicobacter pylor.	220	17	2.1	1168	22	AA163563	Human immune/haema
148	17	2.1	514	22	ABT08113	Myb-related transcr	221	17	2.1	1210	21	AA0383294	Human immune/haema
149	17	2.1	535	22	ABT08113	Human immune/haema	222	17	2.1	1217	22	AA0383294	Human immune/haema
150	17	2.1	553	25	ABZ17991	S2 subtraction lib	223	17	2.1	1235	21	AA0383294	Human immune/haema
151	17	2.1	575	25	ABX66178	Helicobacter pylor	224	17	2.1	1236	21	AA0383294	Human immune/haema
152	17	2.1	580	23	ABV32625	Human prostate exp	225	17	2.1	1236	24	ABT08114	Myb-related transcr
153	17	2.1	580	23	ABV15550	Human prostate exp	226	17	2.1	1305	24	ABZ1859	Human prostate exp
154	17	2.1	580	25	ABZ17775	S2 subtraction lib	227	17	2.1	1333	21	AA035138	Arabidopsis thalia

228	17	2.1	1358	21	AAC43404	Arabidopsis thalia	301	17	2.1	6288	22	AA544629	Fission yeast full
229	17	2.1	1439	21	AAC42906	Arabidopsis thalia	302	17	2.1	6535	22	AA546466	Tumour suppressor
230	17	2.1	1445	22	AAD05792	Arabidopsis thalia	303	17	2.1	6577	22	ABK83688	Human cDNA differe
231	17	2.1	1474	25	AA556036	CMA encoding huma	304	17	2.1	6577	24	ABN95589	Gene #2187 used to
232	17	2.1	1476	19	AAV17815	Homo sapiens don-1	305	17	2.1	6679	21	AA5229170	Partial sequence o
233	17	2.1	1504	21	AACT7864	Human cancer assoc	306	17	2.1	6679	22	AAAD17443	Human granulocyte
234	17	2.1	1506	19	AAV16721	Nucleic acid encod	307	17	2.1	6692	23	AA573329	DNA encoding novel
235	17	2.1	1527	25	ABX77637	Differentially exp	308	17	2.1	6753	22	AAAD17447	Human interferon a
236	17	2.1	1560	21	AAZ96360	A. thaliana gene 1	309	17	2.1	6944	24	ABK34027	Human DNA for stag
237	17	2.1	1576	24	ABZ15877	Arabidopsis thalia	310	17	2.1	6971	24	ABL333237	Human immune syste
238	17	2.1	1584	21	AAD00015	Scorpion protein d	311	17	2.1	7110	22	AA5466489	Tumour suppressor
239	17	2.1	1588	22	AAH77907	Nucleotide sequenc	312	17	2.1	7110	22	ABL70587	Chemically treated
240	17	2.1	1604	24	AB579021	E. coli CFT073 gen	313	17	2.1	7110	24	ABL33556	Human immune syste
241	17	2.1	1668	21	AAC39339	Arabidopsis thalia	314	17	2.1	7110	24	ABG61290	Human gene regulat
242	17	2.1	1670	21	AAC51452	Arabidopsis thalia	315	17	2.1	7146	22	AAE97963	Human immunoglobul
243	17	2.1	1703	22	AAH89921	Human bone marrow	316	17	2.1	7659	21	AAE20934	Human high affinity
244	17	2.1	1711	21	AAC47230	Arabidopsis thalia	317	17	2.1	7659	21	AAA34812	Human adenosine re
245	17	2.1	1729	22	AAH90034	Human bone marrow	318	17	2.1	11968	22	AAO55656	Human reproductive
246	17	2.1	1767	24	AB568489	Fungal peptide syn	319	17	2.1	12291	22	AAK79265	Human immune/haema
247	17	2.1	1769	21	AAC51449	Arabidopsis thalia	320	17	2.1	12543	22	AAH62772	Shrimp white spot
248	17	2.1	1777	21	AAC42051	Arabidopsis thalia	321	17	2.1	15042	23	ABL22246	Drosophila melanog
249	17	2.1	1845	22	AAH27093	Human acyltransfer	322	17	2.1	15881	24	ABL70143	Chemically treated
250	17	2.1	1884	19	AAV17814	Homo sapiens don-1	323	17	2.1	15881	24	ABL32260	Human immune syste
251	17	2.1	2000	24	ABZ16609	Arabidopsis thalia	324	17	2.1	15881	24	ABG61069	Human gene regulat
252	17	2.1	2002	22	AAH14548	Human cDNA sequenc	325	17	2.1	15881	24	ABK31182	Signal transductio
253	17	2.1	2015	22	AA526152	Human cDNA encodin	326	17	2.1	16373	24	ABJ32619	Human immune syste
254	17	2.1	2015	25	ABX73493	Human novel polynu	327	17	2.1	16373	24	ABD28383	Human chemically t
255	17	2.1	2110	22	AAC88085	Human FLEXHR-16 nu	328	17	2.1	17421	22	AA545438	Chemically pretrea
256	17	2.1	2142	24	ABK63660	Rat sequence diffe	329	17	2.1	17421	24	ABK28181	DNA transcription
257	17	2.1	2154	23	ABV25685	Human prostate exp	330	17	2.1	19634	25	ABZ10162	Haematopoietic cel
258	17	2.1	2266	25	AB556045	cDNA encoding huma	331	17	2.1	20530	22	AA532913	Human genomic DNA
259	17	2.1	2268	19	AAV17816	Homo sapiens don-1	332	17	2.1	20530	22	AAO55564	Human reproductive
260	17	2.1	2297	24	ABK12574	DNA encoding LOV k	333	17	2.1	20625	23	ABL226394	Drosophila melanog
261	17	2.1	2391	24	AAZ29173	Candida albicans e	334	17	2.1	21742	21	AAE20938	Human high affinity
262	17	2.1	2647	24	AAD35218	Human TRNFR-3 cDNA	335	17	2.1	21742	21	AAZ27029	Human adenosine re
263	17	2.1	2655	23	ABL16446	Drosophila melanog	336	17	2.1	23071	20	AAK21435	Human CHD1 genomic
264	17	2.1	2718	24	ABN91482	Staphylococcus epi	337	17	2.1	25955	22	ABBA19852	Human nervous syst
265	17	2.1	2734	12	AAQ14293	Modified Protein C	338	17	2.1	34688	24	ABG67059	Human angiogenesis
266	17	2.1	2834	21	AAZ29173	Targeting sequence	339	17	2.1	42115	24	AB568452	C. heterotrophus
267	17	2.1	2883	16	AAO84591	Human CFBP gene.	340	17	2.1	54000	25	AAD52261	Human interferon g
268	17	2.1	2889	23	ABL15734	Drosophila melanog	341	17	2.1	56686	23	ABL18448	Drosophila melanog
269	17	2.1	2927	22	AA502413	Human secreted pro	342	17	2.1	117609	21	AAE21435	Human receptor-rel
270	17	2.1	2933	20	ABZ27028	Human CHD1 genomic	343	17	2.1	237961	24	ABG80552	Human Carlon gene
271	17	2.1	2937	24	AAQ73260	Human ribosomal pr	344	17	2.1	240825	22	AAE24497	Human PG-3 gene
272	17	2.1	3087	21	AAC46767	Arabidopsis thalia	345	17	2.1	240825	24	ABO81802	Human PG-3 gene SB
273	17	2.1	3178	19	AAV42018	Glycerol-3-phospha	346	17	2.1	305107	22	AAH62689	Shrimp white spot
274	17	2.1	3178	19	AAV35739	Klebsiella pneumon	347	17	2.1	513445	22	AAI61373	Soybean j18013 reg
275	17	2.1	3178	19	AAV30824	Mitochondrial glyc	348	17	2.1	640661	24	ABA92787	Buchnera sp. genom
276	17	2.1	3178	20	AAH80613	GUT2 gene. Saccha	349	17	2.1	640661	24	ABA92787	Buchnera sp. genom
277	17	2.1	3337	22	AAH54291	S. epidermidis gen	350	17	2.1	786431	24	ABQ74964	Human kinase prote
278	17	2.1	3555	22	AAH54848	S. epidermidis gen	351	17	2.1	1163020	24	ABQ67197	Listeria innocua
279	17	2.1	3827	22	AAH54867	S. epidermidis gen	352	17	2.1	1503841	24	ABT00010	Human neutregulin
280	17	2.1	3847	24	ABK87764	Human cDNA encodin	353	17	2.1	1503841	24	ABT01503	Human neutregulin
281	17	2.1	3943	22	AAK79264	Human immune/haema	354	17	2.1	1503900	22	AAK95240	Human neutregulin
282	17	2.1	4098	24	ABO54497	Human ovarian anti	355	17	2.1	1503900	22	AAK96673	Human neutregulin
283	17	2.1	4110	25	ABZ10055	Haematopoietic cel	356	17	2.1	1664976	19	AAV21209	Methanococcus jan
284	17	2.1	4110	25	ABZ10201	Haematopoietic cel	357	17	2.1	1664976	19	AAV21209	Methanococcus jan
285	17	2.1	4450	22	AAK51608	Human polynucleoti	358	17	2.1	2365589	24	ABA90521	Genomic sequence
286	17	2.1	4470	22	ABAO9174	Human secreted pro	359	17	2.1	3011208	24	ABO65245	Listeria innocua
287	17	2.1	4470	22	AAK52592	Human polynucleoti	360	17	1.9	27	19	AAK09196	Human diallelic pro
288	17	2.1	4487	12	AAQ14290	Protein C gene fto	361	17	1.9	48	21	AAA30553	C. tropicalis Cys
289	17	2.1	4514	20	AAV91103	Group B Streptococ	362	17	1.9	48	24	ABK31872	Candida tropicalis
290	17	2.1	4515	18	AAV74514	Staphylococcus aur	363	17	1.9	80	21	AA557559	Nucleotide sequenc
291	17	2.1	4779	22	AAI61042	Human polynucleoti	364	17	1.9	108	21	AAI14377	Human secreted pro
292	17	2.1	4779	22	AAI61042	Human polynucleoti	365	17	1.9	108	21	AAI10483	Human diallelic pro
293	17	2.1	5126	24	ABL70493	Chemically treated	366	17	1.9	132	19	AAK10483	Human diallelic pro
294	17	2.1	5304	24	ABL56518	Nucleotide sequenc	367	17	1.9	147	25	ABK49285	Bovine EST associat
295	17	2.1	5470	22	AA526619	Human genomic DNA	368	17	1.9	160	22	ABAA47532	Human breast cell
296	17	2.1	5470	25	ABX73698	Human novel polynu	369	17	1.9	160	22	ABAA65424	Human foetal liver
297	17	2.1	6136	24	ABK40024	Human chemically p	370	17	1.9	160	22	ABK32521	Probe #10987 for g
298	17	2.1	6136	24	ABL70342	Chemically treated	371	17	1.9	160	22	AAK13832	Human brain expres
299	17	2.1	6136	24	AA561297	Human gene regulat	372	17	1.9	160	22	AAK39576	Human bone marrow
300	17	2.1	6235	21	AAZ29169	Human G-CSF genom	373	17	1.9	160	22	AAI20386	Probe #10319 for g
	17	2.1	6235	20	AAK08523	NBP46 (root lectin	374	17	1.9	160	22	AAI45590	Probe #14746 used



C 374	16	1.9	160	22	AA106088	Probe #6079 used t
C 375	16	1.9	160	23	AB39156	Human liver single
C 376	16	1.9	160	24	AB313667	Human genome-deriv
C 377	16	1.9	168	24	ABK72926	Bacillus lichenifo
C 378	16	1.9	184	19	AAK12709	Human biallelic po
C 379	16	1.9	184	19	AAK12710	Human biallelic po
C 380	16	1.9	186	24	ABN69903	Streptococcus poly
C 381	16	1.9	208	23	AA548818	Streptococcus aur
C 382	16	1.9	212	22	AAH71314	Human cervical can
C 383	16	1.9	219	22	AAH7895	Nucleotide sequenc
C 384	16	1.9	225	22	AAK59777	Human immune/haema
C 385	16	1.9	236	25	ABX55724	Bovine EST associa
C 386	16	1.9	248	22	AA169944	Probe m12. Synth
C 387	16	1.9	255	25	ABX27009	Human GDP-mannose
C 388	16	1.9	255	21	AA50098	Arabidopsis herbic
C 389	16	1.9	266	25	ABX3880	Corn ear-derived p
C 390	16	1.9	269	22	ABAI4427	Human nervous syst
C 391	16	1.9	270	22	AAH69889	Human cervical can
C 392	16	1.9	273	18	AAE67757	H. pylori transmem
C 393	16	1.9	273	18	AAE77437	Human ORF2436 CDNA
C 394	16	1.9	279	24	ABN77489	CDNA encoding lung
C 395	16	1.9	292	24	ABK39733	Human lung squamu
C 396	16	1.9	292	25	ACA12062	Lung cancer therap
C 397	16	1.9	292	25	ACA03248	Human ORFX polynu
C 398	16	1.9	294	24	ABN6419	Human gene express
C 399	16	1.9	300	20	AAZ14439	Nucleotide sequenc
C 400	16	1.9	303	22	AAH77893	Human colon cancer
C 401	16	1.9	305	24	ABO56637	Human breast cance
C 402	16	1.9	308	22	AAI19853	Corn ear-derived p
C 403	16	1.9	312	25	ABX88664	DNA encoding novel
C 404	16	1.9	319	23	AA580020	Yeast G-protein co
C 405	16	1.9	319	24	ABK95582	Human ovarian can
C 406	16	1.9	332	24	ABIA1286	S2 subtraction lib
C 407	16	1.9	336	25	ABZ17747	Nucleotide sequenc
C 408	16	1.9	342	22	AAH77891	CDNA encoding huma
C 409	16	1.9	347	25	ACA04740	Human GDP-mannose
C 410	16	1.9	359	25	ABX21401	Cat flea hindgut a
C 411	16	1.9	360	21	AAAC9332	Human secreted pro
C 412	16	1.9	362	21	AAAC08238	Human ORFX polynu
C 413	16	1.9	363	24	ABN18147	Human immune/haema
C 414	16	1.9	364	22	AAK60520	Human immune/haema
C 415	16	1.9	367	25	ABX19802	Human immune/haema
C 416	16	1.9	369	22	AAK61446	Human immune/haema
C 417	16	1.9	375	22	AAK59227	Human immune/haema
C 418	16	1.9	376	22	AAK38337	Novel human diagno
C 419	16	1.9	386	22	AAK74768	Human immune/haema
C 420	16	1.9	388	22	AAK59284	Novel human diagno
C 421	16	1.9	388	22	ABN27072	Human ORFX polynu
C 422	16	1.9	397	22	AAI90589	Human polynucleoti
C 423	16	1.9	398	22	AAI35062	Human musculoskele
C 424	16	1.9	398	22	AAK91526	DNA encoding novel
C 425	16	1.9	398	22	ABX58050	CDNA encoding nove
C 426	16	1.9	400	25	ABX62685	Arabidopsis thalia
C 427	16	1.9	402	21	AAAC24310	Human secreted pro
C 428	16	1.9	402	25	ABX62649	Arabidopsis thalia
C 429	16	1.9	407	22	AAH87931	Peppermint plant o
C 430	16	1.9	408	24	ABZ13643	Arabidopsis thalia
C 431	16	1.9	410	25	ABX48491	Bovine EST associa
C 432	16	1.9	412	22	AAI80754	Human polynucleoti
C 433	16	1.9	412	24	AAH235592	Human gene express
C 434	16	1.9	429	22	AAH52572	S. epidermidis ope
C 435	16	1.9	429	22	AAH53583	S. epidermidis ope
C 436	16	1.9	429	22	AAH53583	Staphylococcus epi
C 437	16	1.9	434	24	ABX61458	Arabidopsis thalia
C 438	16	1.9	440	22	AAK80133	Human immune/haema
C 439	16	1.9	440	22	AAK80134	Human immune/haema
C 440	16	1.9	440	22	AAK80134	DNA encoding novel
C 441	16	1.9	467	22	ABK42393	Human breast cell
C 442	16	1.9	467	22	ABK52821	Human fetal liver
C 443	16	1.9	467	22	ABK22605	Probe #1071 for ge
C 444	16	1.9	467	22	AAK53646	Human cardiovascu
C 445	16	1.9	467	22	AAK01074	Human brain expres
C 446	16	1.9	467	22	AAK26532	Human bone marrow
C 447	16	1.9	467	22	AAI11162	Probe #1095 for ge
C 448	16	1.9	467	22	AAI32425	Probe #1111 used t
C 449	16	1.9	467	22	AAI01080	Probe #1071 used t
C 450	16	1.9	467	23	ABK26124	Human liver single
C 451	16	1.9	467	24	ABK01123	Human genome-deriv
C 452	16	1.9	468	24	ABN66343	Streptococcus poly
C 453	16	1.9	472	22	ABK54736	Human foetal liver
C 454	16	1.9	472	22	AAK76141	Human immune/haema
C 455	16	1.9	472	22	AAI13035	Probe #2968 for ge
C 456	16	1.9	475	21	AAK37744	Arabidopsis thalia
C 457	16	1.9	477	22	AAK79527	Human immune/haema
C 458	16	1.9	503	24	ABN65087	Human cancer relat
C 459	16	1.9	507	18	AAI48200	ITS2 derived from
C 460	16	1.9	508	25	ABT21633	Breast cancer mark
C 461	16	1.9	512	23	AA585013	DNA encoding novel
C 462	16	1.9	514	24	ABO84754	Chlamydia psittaci
C 463	16	1.9	514	24	ABO78012	Chlamydia polynucl
C 464	16	1.9	514	24	ABN26345	Human ORFX polynucl
C 465	16	1.9	514	25	ABX99179	Human ORFX polynucl
C 466	16	1.9	521	22	AAH97889	C. psittaci genom
C 467	16	1.9	527	21	AAO89785	Murine 7-transmemb
C 468	16	1.9	542	21	AAO94996	Cat flea hindgut a
C 469	16	1.9	547	24	ABK69007	Novel murine polyn
C 470	16	1.9	547	24	ABK69351	Novel murine polyn
C 471	16	1.9	551	22	AAI18846	Human ovarian tumo
C 472	16	1.9	559	24	ABO45132	Oligonucleotide fo
C 473	16	1.9	559	24	ABO45133	Oligonucleotide fo
C 474	16	1.9	560	22	ABK64331	Human foetal liver
C 475	16	1.9	560	22	AAK12796	Human brain expres
C 476	16	1.9	560	22	AAK38521	Human bone marrow
C 477	16	1.9	573	21	AAK69563	Streptococcus poly
C 478	16	1.9	575	21	AAE14292	Aspergillus oryzae
C 479	16	1.9	579	24	ABN64027	Human cancer relat
C 480	16	1.9	579	24	ABK62802	Rat sequence diffe
C 481	16	1.9	589	23	ABK37336	Human liver single
C 482	16	1.9	594	22	ABK63387	Human foetal liver
C 483	16	1.9	594	22	ABK63387	Probe #9057 for ge
C 484	16	1.9	594	22	AAK11923	Human brain expres
C 485	16	1.9	594	22	AAK37628	Human bone marrow
C 486	16	1.9	594	22	AAI18385	Probe #8318 for ge
C 487	16	1.9	594	22	AAI43499	Probe #12185 used
C 488	16	1.9	594	23	ABK37282	Human liver single
C 489	16	1.9	594	23	ABK31614	Human genome-deriv
C 490	16	1.9	595	23	ABK58009	Human liver single
C 491	16	1.9	599	24	ABK75659	Bacillus lichenifo
C 492	16	1.9	603	24	ABN62388	Human cancer relat
C 493	16	1.9	603	24	ABN65469	Human cancer relat
C 494	16	1.9	610	23	ABV51719	Human prostate exp
C 495	16	1.9	610	24	ABO66298	Arabidopsis thalia
C 496	16	1.9	617	25	ABZ55497	Aspergillus oryzae
C 497	16	1.9	618	21	AAI16213	Human colon cancer
C 498	16	1.9	625	23	ABK95581	Yeast G-protein co
C 499	16	1.9	627	23	AAK72593	DNA encoding novel
C 500	16	1.9	633	24	ABK53333	Human eosinophil-m
C 501	16	1.9	637	22	AAH07972	Human CDNA clone (
C 502	16	1.9	637	24	ABO89567	Human prostate exp
C 503	16	1.9	646	24	ABL99808	Human secretory po
C 504	16	1.9	648	23	AAK64261	DNA encoding novel
C 505	16	1.9	648	23	AAK78088	DNA encoding novel
C 506	16	1.9	648	23	AAK89070	DNA encoding novel
C 507	16	1.9	650	22	AAH05946	Human CDNA clone (
C 508	16	1.9	657	24	ABO33196	Oligonucleotide fo
C 509	16	1.9	657	24	ABO33197	Oligonucleotide fo
C 510	16	1.9	657	24	ABN68736	Streptococcus poly
C 511	16	1.9	661	24	ABO61148	HSPC10 encoding s
C 512	16	1.9	665	22	AAI34740	Human musculoskele
C 513	16	1.9	665	25	ABK57728	CDNA encoding nove
C 514	16	1.9	670	22	ABK07951	Human ovarian and
C 515	16	1.9	670	22	AAI03861	Human reproductive
C 516	16	1.9	674	22	ABZ16948	Arabidopsis thalia
C 517	16	1.9	681	21	AAK08469	Fusarium venenatum
C 518	16	1.9	684	21	AAK13753	Aspergillus oryzae
C 519	16	1.9	684	22	AAK80663	Human immune/haema



520	16	1.9	690	20	AAK61726	B. burgdorferi ant	593	16	1.9	1143	19	AAV29233	Nucleotide sequenc
521	16	1.9	692	24	ABST77080	Frog embryonic gen	594	16	1.9	1143	20	AAZ23325	Rice DHPS cDNA.
522	16	1.9	693	24	ABQ76541	C. albicans BAX-as	595	16	1.9	1143	20	AAZ23328	Rice DHPS DNA. O
523	16	1.9	713	23	ABLT2567	Drosophila melanog	596	16	1.9	1143	20	AAZ23329	Rice DHPS DNA #2.
524	16	1.9	714	22	AAH04016	Human cDNA clone (	597	16	1.9	1143	21	AAZ15822	Human prostate can
525	16	1.9	714	22	AAH28866	Human prostate can	598	16	1.9	1143	21	AAZ15822	Human prostate can
526	16	1.9	715	21	AACT10171	Human secreted pro	599	16	1.9	1179	22	AAJ36449	Human DNA sequence
527	16	1.9	732	21	ABZ12793	Arabidopsis thalia	600	16	1.9	1179	22	AAJ36450	Human musculocele
528	16	1.9	737	22	AAK61665	Human immune/haema	601	16	1.9	1179	24	ABZ12239	Arabidopsis thalia
529	16	1.9	738	22	AAK80664	Human immune/haema	602	16	1.9	1179	25	ABX59437	CDNA encoding nove
530	16	1.9	738	22	AAK91070	Human digestive sy	603	16	1.9	1179	25	ABX59438	CDNA encoding nove
531	16	1.9	738	22	AAK91070	Human liver associ	604	16	1.9	1193	23	AAH78692	CDNA encoding nove
532	16	1.9	738	24	ABN90460	Human liver antige	605	16	1.9	1193	24	ABO50112	Oligonucleotide fo
533	16	1.9	746	21	AAAC59002	Human secreted pro	606	16	1.9	1194	24	ABO50113	Oligonucleotide fo
534	16	1.9	747	21	ABZ14829	Arabidopsis thalia	607	16	1.9	1195	18	AAH83816	DNA encoding 1 unk
535	16	1.9	750	24	ABZ11725	Human polynucleoti	608	16	1.9	1195	22	AAH74215	DNA encoding eviro
536	16	1.9	762	20	AAZ16628	Human gene express	609	16	1.9	1220	22	ABL64423	Stomach cancer rel
537	16	1.9	763	23	AAH83577	DNA encoding novel	610	16	1.9	1252	21	AAH33127	Arabidopsis thalia
538	16	1.9	763	23	AAH83579	DNA encoding novel	611	16	1.9	1252	24	ABH89865	Human cDNA encodin
539	16	1.9	766	20	AAK61725	B. burgdorferi ant	612	16	1.9	1258	25	ABX05163	Human polynucleoti
540	16	1.9	777	21	AACT7422	Human ORFX ORP297	613	16	1.9	1259	25	ABX05163	Human novel polynu
541	16	1.9	789	22	AAI96433	Human neuroblastom	614	16	1.9	1275	24	ABH92650	Staphylococcus epi
542	16	1.9	800	24	ABST77354	Frog embryonic gen	615	16	1.9	1288	23	AAH72591	DNA encoding novel
543	16	1.9	800	24	ABST77359	Frog embryonic gen	616	16	1.9	1297	24	ABQ54312	Human ovarian anti
544	16	1.9	815	19	AAV04612	Flea serine protea	617	16	1.9	1301	21	AAH39096	Arabidopsis thalia
545	16	1.9	815	22	AAH08679	Flea serine protea	618	16	1.9	1302	23	AAH52054	Staphylococcus aur
546	16	1.9	827	22	AAH08674	Human cDNA clone (	619	16	1.9	1302	23	AAH52054	Staphylococcus aur
547	16	1.9	837	22	AAH53176	S. epidermidis ope	620	16	1.9	1329	23	AAH74685	DNA encoding novel
548	16	1.9	843	17	AAI31065	Human deoxycytidin	621	16	1.9	1332	24	ABZ13948	Arabidopsis thalia
549	16	1.9	848	22	AAI00646	Human reproductive	622	16	1.9	1332	24	AAH54472	Arabidopsis thalia
550	16	1.9	848	24	ABQ23232	Oligonucleotide fo	623	16	1.9	1338	21	AAH53148	Arabidopsis thalia
551	16	1.9	848	24	ABQ23233	Oligonucleotide fo	624	16	1.9	1339	12	AAO14294	Truncated Protein
552	16	1.9	850	20	AAH39974	Prostate cancer as	625	16	1.9	1340	22	AAH45233	Arabidopsis thalia
553	16	1.9	855	21	AACT7424	Human ORFX ORP297	626	16	1.9	1340	22	AAH60607	Human polynucleoti
554	16	1.9	856	21	AACT9668	Arabidopsis thalia	627	16	1.9	1355	21	AAH69089	Human secreted pro
555	16	1.9	868	24	ABH64929	Maize ribonuclease	628	16	1.9	1357	21	AAH34088	Arabidopsis thalia
556	16	1.9	879	21	AACT4537	Arabidopsis thalia	629	16	1.9	1415	21	AAH34689	Arabidopsis thalia
557	16	1.9	884	25	ABH76193	Lung cancer-associ	630	16	1.9	1422	25	AAH37231	CDNA encoding nove
558	16	1.9	907	20	AAZ53005	Human prostate tum	631	16	1.9	1422	25	ABH60219	CDNA encoding nove
559	16	1.9	925	24	ABH65341	Arabidopsis thalia	632	16	1.9	1428	21	AAH33940	Arabidopsis thalia
560	16	1.9	933	22	AAH85150	Nucleotide sequenc	633	16	1.9	1438	22	AAH18115	Human cDNA sequenc
561	16	1.9	939	24	ABH67934	Streptococcus poly	634	16	1.9	1439	23	AAH85014	DNA encoding novel
562	16	1.9	943	24	ABH98584	Arabidopsis thalia	635	16	1.9	1443	21	AAH47447	Arabidopsis thalia
563	16	1.9	944	21	AAH40600	Arabidopsis thalia	636	16	1.9	1445	21	AAH39105	Arabidopsis thalia
564	16	1.9	952	21	AAH47445	Arabidopsis thalia	637	16	1.9	1453	21	AAH33607	Arabidopsis thalia
565	16	1.9	953	21	AAH35444	Arabidopsis thalia	638	16	1.9	1470	22	AAH53837	S. epidermidis ope
566	16	1.9	956	24	ABH14221	A. thaliana Produc	639	16	1.9	1470	22	AAH53892	S. epidermidis ope
567	16	1.9	957	19	AAV52504	Streptococcus pneu	640	16	1.9	1476	21	AAH47228	Arabidopsis thalia
568	16	1.9	958	19	AAV33481	Human disease rela	641	16	1.9	1476	24	ABH04755	Chlamydia psittaci
569	16	1.9	966	22	AAH69953	Human immune/haema	642	16	1.9	1476	24	ABH78013	Chlamydia polynuci
570	16	1.9	983	24	ABH74737	Bovine embryonic g	643	16	1.9	1476	25	ABH99180	C. psittaci genom
571	16	1.9	992	21	AAH42424	Arabidopsis thalia	644	16	1.9	1478	21	AAH42167	Arabidopsis thalia
572	16	1.9	995	18	AAH84194	DNA encoding a Sta	645	16	1.9	1480	21	AAH36611	Arabidopsis thalia
573	16	1.9	1001	21	AAH57822	Arachidonic acid m	646	16	1.9	1496	22	AAH76830	Human promoter bin
574	16	1.9	1003	21	AAH41613	Arabidopsis thalia	647	16	1.9	1508	22	AAH33462	Human colon cancer
575	16	1.9	1003	21	AAH52036	Arabidopsis thalia	648	16	1.9	1511	22	AAH77238	Human immune/haema
576	16	1.9	1004	24	ABH97970	Human secretory po	649	16	1.9	1511	22	AAH77240	Human immune/haema
577	16	1.9	1006	21	AAH34012	Arabidopsis thalia	650	16	1.9	1518	19	AAH16720	Human ovary and
578	16	1.9	1011	21	AAH77428	Human ORFX ORP2983	651	16	1.9	1518	22	AAH77690	Nucleotide sequenc
579	16	1.9	1026	24	ABH23204	Candida albicans e	652	16	1.9	1525	22	ABH07949	Human reproductive
580	16	1.9	1038	24	ABH07946	Human NOV6a protei	653	16	1.9	1525	22	AAH03859	Human reproductive
581	16	1.9	1059	21	AAH43418	Arabidopsis thalia	654	16	1.9	1552	24	ABH15058	A. thaliana ligand
582	16	1.9	1059	21	ABH212339	Arabidopsis thalia	655	16	1.9	1561	25	ABH56034	CDNA encoding muti
583	16	1.9	1080	22	AAH99788	Human protein enco	656	16	1.9	1587	24	ABH79877	Fungal ZBC gene se
584	16	1.9	1089	23	ABH06597	Drosophila melanog	657	16	1.9	1607	19	AAH17813	Mus musculus don-1
585	16	1.9	1104	21	AAH45776	Arabidopsis thalia	658	16	1.9	1633	22	AAH29147	Polynucleotide seq
586	16	1.9	1110	21	AAH41607	Arabidopsis thalia	659	16	1.9	1636	24	ABH77199	Human p21-activate
587	16	1.9	1113	21	AAH43438	Arabidopsis thalia	660	16	1.9	1637	22	AAH37232	Human musculocele
588	16	1.9	1113	21	AAH43438	Arabidopsis thalia	661	16	1.9	1637	22	AAH37232	Human musculocele
589	16	1.9	1122	23	AAH77737	DNA encoding novel	662	16	1.9	1657	25	ABH60220	CDNA encoding nove
590	16	1.9	1131	22	AAH99803	Human protein enco	663	16	1.9	1658	23	AAH68022	DNA encoding novel
591	16	1.9	1134	24	ABH72837	Bacillus lichenifo	664	16	1.9	1659	16	AAH74818	Firefly luciferase
592	16	1.9	1134	24	ABH66151	Streptococcus poly	665	16	1.9	1662	21	AAH40545	Arabidopsis thalia

666	16	1.9	1701	20	AAK21366	Coding sequence fo	739	16	1.9	2546	23	ABL24542	Drosophila melanog
667	16	1.9	1704	22	ABR14402	Human nervous syst	740	16	1.9	2571	22	AA158422	Human polynucleoti
668	16	1.9	1705	16	AAQ79621	Schizosaccharomyce	741	16	1.9	2579	24	AA141594	HIV-1 subtype C 1s
669	16	1.9	1721	25	ABZ67722	Human secreted pro	742	16	1.9	2610	24	ABK51388	cDNA encoding huma
670	16	1.9	1736	22	AA526870	Human cDNA encodin	743	16	1.9	2648	22	AA160208	Human polynucleoti
671	16	1.9	1737	22	AAH64793	Human secreted pro	744	16	1.9	2659	19	AAV34185	Human secreted pro
672	16	1.9	1738	21	AACT4301	Human secreted pro	745	16	1.9	2668	23	AA571214	DNA encoding novel
673	16	1.9	1738	25	ABZ67087	Human secreted pro	746	16	1.9	2671	23	AA582970	DNA encoding novel
674	16	1.9	1739	22	AAH16813	Human cDNA sequenc	747	16	1.9	2681	24	ABK63551	Rat sequence diffe
675	16	1.9	1748	22	AA158821	Human polynucleoti	748	16	1.9	2682	22	AAH77602	Human ubiquitin-pr
676	16	1.9	1757	22	AAH64800	Human secreted pro	749	16	1.9	2687	23	AB105552	Drosophila melanog
677	16	1.9	1773	22	AA521983	Human collagen gen	750	16	1.9	2784	22	ABH18856	Human nervous syst
678	16	1.9	1775	22	AAAC42335	Arabidopsis thalia	751	16	1.9	2812	23	AA582457	DNA encoding novel
679	16	1.9	1803	18	AAAT87923	Rat cerebellum der	752	16	1.9	2812	23	AA593151	Drosophila melanog
680	16	1.9	1809	18	AAAT84169	DNA encoding a Sta	753	16	1.9	2817	23	AB121612	AB121612
681	16	1.9	1815	21	AAFI8138	Lung cancer associ	754	16	1.9	2821	23	AB122566	Drosophila melanog
682	16	1.9	1833	24	ABZ17541	Arabidopsis thalia	755	16	1.9	2826	23	AA551976	Staphylococcus aur
683	16	1.9	1840	23	AA580692	DNA encoding novel	756	16	1.9	2847	23	AA554776	Staphylococcus aur
684	16	1.9	1859	22	AAH46945	Human secreted pro	757	16	1.9	2879	23	AA572595	Drosophila melanog
685	16	1.9	1859	24	AB558481	Human secreted pro	758	16	1.9	2883	23	AB122588	Drosophila melanog
686	16	1.9	1869	22	AAAF1515	Tobacco PRP-amido	759	16	1.9	2910	21	AA552178	DNA encoding H. in
687	16	1.9	1871	24	ABO72513	Human MDTF encodin	760	16	1.9	2928	21	AA552177	Haemophilus influe
688	16	1.9	1888	24	ABV74635	Human nucleotide e	761	16	1.9	2946	22	AAH62796	Shrimp white spot
689	16	1.9	1903	24	ABK84290	Human cDNA differe	762	16	1.9	2961	22	ABX04186	Human mRNA differe
690	16	1.9	1903	24	ABK61902	Colon adenocarcino	763	16	1.9	3001	21	AAH51704	Chromosome 13q11-q
691	16	1.9	1915	23	AA586627	DNA encoding novel	764	16	1.9	3025	23	ABV21114	Human prostate exp
692	16	1.9	1918	21	AACT77195	Human ORFX ORP2750	765	16	1.9	3025	23	ABV26956	Human prostate exp
693	16	1.9	1938	22	ABH18055	Human nervous syst	766	16	1.9	3051	24	ABO99464	Human coding seque
694	16	1.9	1962	22	AAAF1263	Tobacco dihydro-ot	767	16	1.9	3076	19	AAV43674	Receptor type tyro
695	16	1.9	1975	19	AAZ96228	S. pneumoniae deri	768	16	1.9	3089	23	AB106596	Drosophila melanog
696	16	1.9	1975	19	AAV42962	Streptococcus pneu	769	16	1.9	3105	21	AAZ51819	Ruminoecoccus xylan
697	16	1.9	1990	22	AAH64898	Human secreted pro	770	16	1.9	3119	22	AAH54906	S. epidermidis gen
698	16	1.9	2000	24	ABZ15141	Arabidopsis thalia	771	16	1.9	3156	23	AB127600	Drosophila melanog
699	16	1.9	2000	24	ABZ16898	Arabidopsis thalia	772	16	1.9	3167	22	AAH99522	Human protein enco
700	16	1.9	2000	24	ABZ17029	Arabidopsis thalia	773	16	1.9	3201	24	AB232025	Candida albicans e
701	16	1.9	2001	24	ABN67971	Streptococcus poly	774	16	1.9	3233	22	AAH54222	S. epidermidis gen
702	16	1.9	2023	21	AAAC49400	Arabidopsis thalia	775	16	1.9	3234	24	AB232644	Candida albicans e
703	16	1.9	2042	21	AAAC41973	Human NOV6b protei	776	16	1.9	3249	22	AAH54258	S. epidermidis gen
704	16	1.9	2058	24	ABO79947	Pfirefly luciferase	777	16	1.9	3252	23	AB106532	Drosophila melanog
705	16	1.9	2063	16	AAQ74819	Salmonella typhi D	778	16	1.9	3294	22	AA159648	Human polynucleoti
706	16	1.9	2115	22	AA556230	Human cDNA sequenc	779	16	1.9	3294	22	AA159649	Human polynucleoti
707	16	1.9	2175	22	AAH18603	DNA encoding novel	780	16	1.9	3324	23	AA594438	DNA encoding novel
708	16	1.9	2182	23	AA583580	Human encoding novel	781	16	1.9	3381	23	AA551736	Staphylococcus aur
709	16	1.9	2187	24	ABN91966	Staphylococcus epi	782	16	1.9	3434	23	AB127586	Drosophila melanog
710	16	1.9	2193	24	AA521157	Worm cDNA sequenc	783	16	1.9	3441	18	AAAT87922	Rat cerebellum der
711	16	1.9	2193	25	ABX74424	Human cDNA sequenc	784	16	1.9	3477	23	AA554787	Staphylococcus aur
712	16	1.9	2220	24	ABK72811	Bacillus lichenifo	785	16	1.9	3494	22	AAH54286	S. epidermidis gen
713	16	1.9	2253	24	ABZ14355	Arabidopsis thalia	786	16	1.9	3496	24	ABK75331	Bacillus lichenifo
714	16	1.9	2270	22	AA104591	Human reproductive	787	16	1.9	3504	22	AACT87907	Arabidopsis thalia
715	16	1.9	2270	22	AA104693	Human reproductive	788	16	1.9	3511	22	AAH54566	S. epidermidis gen
716	16	1.9	2270	23	AA104693	Human testicular a	789	16	1.9	3579	25	ABX10378	DNA encoding prote
717	16	1.9	2270	23	AB197514	Human testicular a	790	16	1.9	3612	22	AAH54332	S. epidermidis gen
718	16	1.9	2298	22	AA506720	Polynucleotide seq	791	16	1.9	3619	22	AAH53971	S. epidermidis gen
719	16	1.9	2317	21	AAH67992	Sea Island Fblate2	792	16	1.9	3702	24	ABN86068	Human arylglydrolar
720	16	1.9	2324	23	AB127552	Drosophila melanog	793	16	1.9	3738	21	AAH70178	Plasmodium falcipa
721	16	1.9	2369	23	AB126922	Drosophila melanog	794	16	1.9	3741	23	AB105790	Drosophila melanog
722	16	1.9	2387	24	AAAD33776	Mouse MxgB8 (mas-r	795	16	1.9	3830	22	AAH54570	S. epidermidis gen
723	16	1.9	2391	24	AA571212	DNA encoding novel	796	16	1.9	3842	22	AACT87905	Arabidopsis thalia
724	16	1.9	2391	23	AA572592	DNA encoding novel	797	16	1.9	3856	23	AB105792	Drosophila melanog
725	16	1.9	2405	22	ABH13519	Human nervous syst	798	16	1.9	3882	22	ABA20413	Human nervous syst
726	16	1.9	2410	23	AB126976	Drosophila melanog	799	16	1.9	3882	22	ABA20701	Human nervous syst
727	16	1.9	2427	24	AB163303	Breast cancer rela	800	16	1.9	3974	18	AAAT43360	Cotton Fblate2-82A
728	16	1.9	2430	25	ABX08016	S. pneumoniae type	801	16	1.9	3966	23	AA578739	DNA encoding novel
729	16	1.9	2433	21	AAZ91829	Streptococcus pneu	802	16	1.9	4031	24	ABO55072	Human ovarian anti
730	16	1.9	2433	23	AA555901	Streptococcus pneu	803	16	1.9	4158	22	AA103882	Human reproductive
731	16	1.9	2442	25	AB556033	cDNA encoding murt	804	16	1.9	4158	22	AAK68651	Human immune/naema
732	16	1.9	2463	23	AA572590	DNA encoding novel	805	16	1.9	4165	24	ABK83664	Human cDNA differe
733	16	1.9	2467	19	AAV17812	Mus musculus don-1	806	16	1.9	4165	24	AB161758	Colon adenocarcino
734	16	1.9	2476	24	ABK65347	Arabidopsis cDNA e	807	16	1.9	4169	22	AA103883	Human reproductive
735	16	1.9	2485	22	AAH16083	Human cDNA sequenc	808	16	1.9	4169	22	AAK68652	Human immune/naema
736	16	1.9	2494	19	AAV67228	Human doublet enco	809	16	1.9	4315	25	ABO83363	Human NR10.3 splic
737	16	1.9	2530	25	ABX05482	Human novel polynu	810	16	1.9	4316	22	ABA83122	Human HLA-Cw ovatio
738	16	1.9	2537	20	AAV80583	Kidney injury asso	811	16	1.9	4316	24	ABN97218	Gene #3716 used to

C 812	16	1.9	4379	22	AAK82496	Human immune/haema
C 813	16	1.9	4384	25	ABT23200	Human protein modi
C 814	16	1.9	4517	20	AAK90200	Human yesi encodin
C 815	16	1.9	4517	22	AAH28359	Nucleotide sequenc
C 816	16	1.9	4517	23	AAK74489	DNA encoding novel
C 817	16	1.9	4517	24	ABV94313	Breast carcinoma r
C 818	16	1.9	4543	22	AAI59755	Human polynucleoti
C 820	16	1.9	4579	22	AAI59759	Human polynucleoti
C 821	16	1.9	4579	23	ABL07748	Drosophila melanog
C 822	16	1.9	4594	18	AAV74511	Staphylococcus aur
C 823	16	1.9	4599	23	ABL10253	Drosophila melanog
C 824	16	1.9	4607	24	ABK46540	DNA encoding human
C 825	16	1.9	4689	21	AAK45991	Arabidopsis thalia
C 826	16	1.9	4698	22	AAH62806	Shrimp white spot
C 827	16	1.9	4772	24	AAAD26488	Human kinase PKI-
C 828	16	1.9	4808	23	ABK66335	Drosophila melanog
C 829	16	1.9	4918	23	ABL03178	Soybean locus (SYS
C 830	16	1.9	4958	23	ABL19334	Drosophila melanog
C 831	16	1.9	5007	24	ABL13467	Human metastasis a
C 832	16	1.9	5042	23	ABL17666	Drosophila melanog
C 833	16	1.9	5077	21	AAK50096	Arabidopsis thalia
C 834	16	1.9	5128	22	AAI57863	Human polynucleoti
C 835	16	1.9	5139	25	ABV72514	Nucleotide sequenc
C 836	16	1.9	5278	24	ABL92239	Chemically treated
C 837	16	1.9	5283	23	ABL3893	Human immune syste
C 838	16	1.9	5344	22	AAI57862	Human polynucleoti
C 839	16	1.9	5414	21	AAK45028	Arabidopsis thalia
C 840	16	1.9	5443	24	ABL32883	Human immune syste
C 841	16	1.9	5518	18	AAI73870	Cotton fibre promo
C 842	16	1.9	5519	24	ABK76437	CDNA encoding huma
C 843	16	1.9	5547	18	AAI73865	Cotton fibre promo
C 844	16	1.9	5582	24	ABK63552	Rat sequence diffe
C 845	16	1.9	5622	21	AAK6506	DNA construct comp
C 846	16	1.9	5675	24	ABL70246	Chemically treated
C 847	16	1.9	5675	24	ABL32969	Human immune syste
C 848	16	1.9	5675	24	ABL34521	Human metastasis a
C 849	16	1.9	5828	22	ABAI9864	Human nervous syst
C 850	16	1.9	5895	18	AAV74312	Staphylococcus aur
C 851	16	1.9	6005	22	AAK72246	Human immune/haema
C 852	16	1.9	6018	23	ABL06075	Drosophila melanog
C 853	16	1.9	6047	23	ABL09681	Drosophila melanog
C 854	16	1.9	6047	24	ABL33458	Human immune syste
C 855	16	1.9	6047	24	ABL33459	Human immune syste
C 856	16	1.9	6055	23	ABV22568	Drosophila melanog
C 857	16	1.9	6073	23	ABV2269	Human prostate exp
C 858	16	1.9	6073	23	ABV28495	Human prostate exp
C 859	16	1.9	6112	22	AAI36593	Human musculoskele
C 860	16	1.9	6112	25	AAK59581	CDNA encoding nove
C 861	16	1.9	6116	24	ABN80129	Human chemically m
C 862	16	1.9	6156	24	ABL92288	Chemically treated
C 863	16	1.9	6156	24	ABL9359	Human polynucleoti
C 864	16	1.9	6210	19	ABL34215	Human immune syste
C 865	16	1.9	6216	24	AAV06946	Canine herpes viru
C 866	16	1.9	6225	23	ABL20912	Drosophila melanog
C 867	16	1.9	6232	22	AAI67936	Chloroplast gene t
C 868	16	1.9	6234	19	AAV64131	St-H segment of Dr
C 869	16	1.9	6246	22	AAK52408	Human collagen gen
C 870	16	1.9	6246	22	AAK52408	Nucleotide sequenc
C 871	16	1.9	6270	17	AAI45728	Cyclotella cryptic
C 872	16	1.9	6271	22	ABAI5878	Human nervous syst
C 873	16	1.9	6271	22	AAI04938	Human reproductive
C 874	16	1.9	6271	22	AAK5395	Human immune/haema
C 875	16	1.9	6271	23	ABL97832	Human testicular a
C 876	16	1.9	6277	24	AAK46728	Tumour suppressor
C 877	16	1.9	6277	24	ABL34040	Human immune syste
C 878	16	1.9	6352	24	ABL70563	Chemically treated
C 879	16	1.9	6352	24	AAK51235	Human gene regulat
C 880	16	1.9	6352	24	ABK31340	Signal transductio
C 881	16	1.9	6354	23	ABL14680	Drosophila melanog
C 882	16	1.9	6370	24	ABL70568	Chemically treated
C 883	16	1.9	6370	24	ABK31349	Signal transductio
C 884	16	1.9	6609	24	ABN80247	Human chemically m
C 885	16	1.9	6609	24	ABL33883	Human immune syste
C 886	16	1.9	6712	21	AAK78691	Human ORFX ORF2246
C 887	16	1.9	6766	22	AAK78691	Human immune/haema
C 888	16	1.9	6766	23	ABK42991	Genomic sequence #
C 889	16	1.9	6790	17	AAI45727	Cyclotella cryptic
C 890	16	1.9	6954	24	ABL33390	Human immune syste
C 891	16	1.9	6995	16	AAQ85466	Ras oncogene/lacZ
C 892	16	1.9	7027	23	ABL16960	Drosophila melanog
C 893	16	1.9	7282	24	ABN59772	Novel human coding
C 894	16	1.9	7344	23	ABL09158	Drosophila melanog
C 895	16	1.9	7380	24	ABK61453	Human CDNA encodin
C 896	16	1.9	7458	22	ABA08075	Human ovarian and
C 897	16	1.9	7458	22	ABA08075	Human ovarian and
C 898	16	1.9	7458	22	AAI06801	Human reproductive
C 899	16	1.9	7458	22	AAI07394	Human reproductive
C 900	16	1.9	7451	24	ABL33784	Human immune syste
C 901	16	1.9	7491	22	AAK78292	Human immune/haema
C 902	16	1.9	7623	24	ABL34180	Human immune syste
C 903	16	1.9	7636	20	AAI3474	Enterococcus faeca
C 904	16	1.9	7696	24	ABK9269	Enterococcus faeca
C 905	16	1.9	7893	20	AAI3165	Enterococcus faeca
C 906	16	1.9	7893	24	ABK98960	Enterococcus faeca
C 907	16	1.9	8032	22	AAI31372	Inventon related
C 908	16	1.9	8108	21	AAK36930	Nucleotide sequenc
C 909	16	1.9	8192	23	ABL10252	Drosophila melanog
C 910	16	1.9	8387	22	AAK84330	Stealth virus nucl
C 911	16	1.9	8514	22	AAK81530	Human immune/haema
C 912	16	1.9	8592	24	ABL33382	Human immune syste
C 913	16	1.9	8626	16	AAQ98237	Partial rice pyruv
C 914	16	1.9	8697	19	AAV52256	Streptococcus pneu
C 915	16	1.9	8811	23	ABL10800	Drosophila melanog
C 916	16	1.9	8860	23	ABL06074	Drosophila melanog
C 917	16	1.9	9046	16	AAI02792	Transcription fact
C 918	16	1.9	9046	16	AAI02792	Transcription fact
C 919	16	1.9	9046	19	AAI02792	Transcription fact
C 920	16	1.9	9370	23	ABL10080	Human p300 gene SB
C 921	16	1.9	9408	21	AAI70157	Drosophila melanog
C 922	16	1.9	9636	15	AAQ67190	Plasmodium falcipa
C 923	16	1.9	9928	22	ABA19865	Human nervous syst
C 924	16	1.9	10207	15	AAQ55134	Staphylococcus aur
C 925	16	1.9	10207	25	ABK77350	Nucleotide sequenc
C 926	16	1.9	10207	25	AAI51838	Staphylococcus aur
C 927	16	1.9	10556	22	AAI59459	Human polynucleoti
C 928	16	1.9	10558	24	ABK04155	Human mRNA diffe
C 929	16	1.9	10558	24	ABK92097	Human Tumour Endot
C 930	16	1.9	10558	24	ABK76357	Human alpha3 colla
C 931	16	1.9	10558	25	ABK72022	DNA encoding human
C 932	16	1.9	10595	22	AAK46542	Tumour suppressor
C 933	16	1.9	10740	22	AAK464788	Human immune/haema
C 934	16	1.9	10864	22	AAK46262	DNA encoding novel
C 935	16	1.9	11011	21	AAK68252	B. burgdorferi tox
C 936	16	1.9	11115	22	AAI41967	Genomic sequence #
C 937	16	1.9	11115	22	AAI41974	Genomic sequence #
C 938	16	1.9	11115	23	ABK72139	Human ovarian anti
C 939	16	1.9	11115	24	ABK91731	Novel ovarian rela
C 940	16	1.9	11260	22	AAK45315	Chemically pretrea
C 941	16	1.9	11260	24	ABN80039	Human chemically m
C 942	16	1.9	11260	24	ABK28154	DNA transcription
C 943	16	1.9	11287	23	ABL21090	Drosophila melanog
C 944	16	1.9	11492	24	ABK58983	AAV expression in
C 945	16	1.9	12175	24	ABL33581	Ryegrass cinamoyl
C 946	16	1.9	12225	23	ABL08714	Drosophila melanog
C 947	16	1.9	12541	22	AAI35819	Human musculoskele
C 948	16	1.9	12541	25	ABK58807	CDNA encoding nove
C 949	16	1.9	13086	18	AAV74327	Staphylococcus aur
C 950	16	1.9	13125	24	ABL70284	Chemically treated
C 951	16	1.9	13125	24	ABL33227	Human immune syste
C 952	16	1.9	13125	24	ABL34557	Human metastasis a
C 953	16	1.9	14518	22	AAK67185	Human immune/haema
C 954	16	1.9	14518	22	AAK67186	Human immune/haema
C 955	16	1.9	14518	22	AAK67187	Human immune/haema
C 956	16	1.9	15734	22	AAK36422	Human cardiovascular
C 957	16	1.9	17211	24	ABL32655	Human immune syste



```
OY 661 GAAATATCACCACTTAATAATTTGTAAGTAAGTCCATGCAACAGGACAAATTAAGT 720
DB 661 GAAATATCACCACTTAATAATTTGTAAGTAAGTCCATGCAACAGGACAAATTAAGT 720
OY 721 CATGAAATTTGGGGTGAATTTTCTTAATTTACACCCATGCAACAGAGTACAAAT 780
DB 721 CATGAAATTTGGGGTGAATTTTCTTAATTTACACCCATGCAACAGAGTACAAAT 780
OY 781 GATGATTTTCTGCTGAATTTGACTTAAGATCTACTGATTTAA 825
DB 781 GATGATTTTCTGCTGAATTTGACTTAAGATCTACTGATTTAA 825

RESULT 2
ABQ73048
ID ABQ73048 standard; cDNA; 1012 BP.
AC ABQ73048;
XX
XX 24-SEP-2002 (first entry)
DT
DE Tomato anthocyanin 1 (ANT1) genomic fragment SEQ ID NO:4.
XX
XX Tomato; ANT1; anthocyanin 1; Lycopersicon esculentum; plant; colour;
KM gene; ds.
XX
XX Lycopersicon esculentum.
OS
XX MO200255658-A2.
XX
XX 18-JUL-2002.
PD
PF 29-OCT-2001; 2001MO-US50638.
XX
XX 30-OCT-2000; 2000US-244685P.
PR
XX (EXEL-) EXELIXIS PLANT SCI INC.
XX
XX Connors K, Mathews HV, Liu A;
PI
XX WPI; 2002-557819/59.
XX
XX New isolated polynucleotide derived from tomato, useful for producing
PT an Anthocyanin 1 phenotype in plants, particularly for modifying e.g.
PT leaf color, flower color or fruit color in plants
XX
XX Example 3; Page 46-47; 53pp; English.
XX
XX The present invention describes tomato anthocyanin 1 (ANT1). The ANT1
XX polynucleotide can be used for modifying the expression of a native
XX plant gene, particularly for producing an anthocyanin 1 phenotype in
XX plants, which is responsible for many red and blue colours in plants.
XX The polynucleotide is useful for modifying e.g. leaf colour, flower
XX colour or fruit colour in plants. The present sequence represents an
XX ANT1 genomic DNA fragment which is used in an example from the present
XX invention.
XX
XX Sequence 1012 BP; 364 A; 138 C; 201 G; 309 T; 0 other;
SQ

Query Match: 68.7%; Score 567; DB 24; Length 1012;
Best Local Similarity 100.0%; Pred. No. 3.6e-273;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 259 AGATGTCACCTATTTGCTGTAGACTTCCGGAGAGACAGTAACGATGGAAGAACTAT 318
DB 446 AGATGTCACCTATTTGCTGTAGACTTCCGGAGAGACAGTAACGATGGAAGAACTAT 505
OY 319 TGAACACTATCTTCTTAAGGAATTAATACTAATAATTTGTTCTCGGAAAGATT 378
DB 506 TGAACACTATCTTCTTAAGGAATTAATACTAATAATTTGTTCTCGGAAAGATT 565
OY 379 AACATTAAGTGTGAGAAATTAAGTAACTTAAGTAAATTAACCTCAACGACGAG 438
PI
```

```
DB 566 AACATTAAGTGTGAGAAATTAAGTAACTTAAGTAAATTAACCTCAACGACGAG 625
OY 439 TATTTCTCAAGCAATTAAGGAATTTTACAAACATTAATGTAATTTTGGACGAGGAA 498
DB 626 TATTTCTCAAGCAATTAAGGAATTTTACAAACATTAATGTAATTTTGGACGAGGAA 685
OY 499 CATTCGAGGAATTAATTAAGTGAAGAAACAACTCCAGATGATGAGCAACGTAAGAT 558
DB 686 CATTCGAGGAATTAATTAAGTGAAGAAACAACTCCAGATGATGAGCAACGTAAGAT 745
OY 559 CCATGGTGAATTAATTTACTGGAATAATTCGATGACGATTTGCAAGAGAGGTT 618
DB 746 CCATGGTGAATTAATTTACTGGAATAATTCGATGACGATTTGCAAGAGAGGTT 805
OY 619 GTAATTAATTAAGAAACAACTAAGAGTTTGTATGATGAAGAAATATGACCACTTA 678
DB 806 GTAATTAATTAAGAAACAACTAAGAGTTTGTATGATGAAGAAATATGACCACTTA 865
OY 679 AATATTGGTGAAGTAAGTCCATGCAACAGCAAAATTAAGTCATGAATAATGGGGTGA 738
DB 866 AATATTGGTGAAGTAAGTCCATGCAACAGCAAAATTAAGTCATGAATAATGGGGTGA 925
OY 739 TTTTCTTAAATTTACCAATGCAACAGAGTAACAAATGATGATTTTCTGCTGAA 798
DB 926 TTTTCTTAAATTTACCAATGCAACAGAGTAACAAATGATGATTTTCTGCTGAA 985
OY 799 ATTGACTTAATGGAATCTACTGATTTAA 825
DB 986 ATTGACTTAATGGAATCTACTGATTTAA 1012

RESULT 3
ABK65274
ID ABK65274 standard; cDNA; 1300 BP.
AC ABK65274;
XX
XX 02-JUL-2002 (first entry)
DT
DE Arabidopsis cDNA encoding a transcription factor #126.
XX
XX Plant; ss; gene; transcription factor; transgenic;
XX agriculture; metabolic chemical; environmental stress; drought;
XX microbial disease resistance; herbicide resistance; seed yield;
XX fruit yield; growth rate; leaf senescence; flower senescence.
XX
XX Arabidopsis thaliana.
OS
XX
XX MO200215675-A1.
XX
XX 28-FEB-2002.
PD
XX 22-AUG-2001; 2001MO-US26189.
XX
XX 22-AUG-2000; 2000US-227439P.
PR 16-NOV-2000; 2000US-0713994.
PR 16-APR-2001; 2001US-0837944.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX (PILG/) PILGRIM M.
XX (CREE/) CREELMAN R.
XX (DUBE/) DUBELL A J.
XX (HEAR/) HEARD J.
XX (JIAN/) JIANG C.
XX (KEDD/) KEDDIE J.
XX (ADAM/) ADAM L.
XX (RATC/) RATCLIFF O.
XX (REUB/) REUBER J L.
XX (RIEC/) RIECHMANN J L.
XX (YUGG/) YU G.
XX (PINE/) PINEDA O.
PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;
```

PI Adam L, Ratcliffe O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
XX WPI: 2002-292022/33.  
DR P-PSDB; AAU93088.  
XX  
PT An isolated or recombinant polynucleotide used to produce a transgenic  
PT plant -  
PS Claim 4; Page 523-525; 941pp; English.  
XX  
XX The invention relates to 1 of 232 isolated or recombinant polynucleotides  
CC encoding an Arabidopsis thaliana transcription factor, their variants,  
CC complements, fragments, or related polynucleotide with 31% to 95%  
CC sequence identity, where the plant possesses an altered trait as compared  
CC to a wild-type or reference plant, or the plant exhibits an altered  
CC phenotype as compared to a wild-type or reference plant, or the plant  
CC exhibits ectopic expression or altered expression of one or more genes  
CC associated with a plant trait as compared to a wild plant. Also included  
CC are a transgenic plant comprising the polynucleotides, a computer  
CC readable medium having stored sequence information, and identifying a  
CC homologue sequence from a database comprising a plurality of known plant  
CC sequences comprising inputting sequence information selected from one of  
CC 464 fully defined sequences given in the specification. The isolated or  
CC recombinant polynucleotide is used for producing a plant having a  
CC modified trait, the method comprising selecting a polynucleotide that  
CC encodes a polypeptide or an antisense nucleic acid, inserting the  
CC polynucleotide or antisense nucleic acid into an expression vector,  
CC introducing the vector into a plant or a cell of a plant to overexpress  
CC the polypeptide or antisense nucleic acid, thereby producing a modified  
CC plant, and selecting for a modified trait (e.g. increased  
CC production of agriculturally useful proteins or metabolic chemicals,  
CC pest tolerance, environmentally stress response (e.g. drought), microbial  
CC disease resistance, herbicide resistance, seed and fruit yield, growth  
CC rate, leaf and flower senescence and many other traits listed in the  
CC specification). The present sequence is one of the 232 polynucleotides  
CC encoding an A. thaliana transcription factor.  
XX  
SQ Sequence 1300 BP; 392 A; 210 C; 350 G; 348 T; 0 other;

Query Match 2.8%; Score 23; DB 24; Length 1300;  
Best Local Similarity 100.0%; Pred. No. 0.33;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 253 GGCAACAGATGTCCTATTGC 275  
DB 423 GGCAACAGATGTCCTATTGC 445

RESULT 4  
AAA01109/c  
ID AAA01109 standard; cDNA; 300 BP.  
XX  
AC AAA01109;  
XX  
DT 19-MAY-2000 (first entry)  
XX  
DE Human colon cancer cell line polynucleotide sequence SEQ ID NO.1100.  
XX  
XX Human; colon cancer; tumour; diagnosis; gene expression product;  
KW probe; detection; cancerous state; metastasis; identification;  
KW breast cancer; oestrogen receptor-positive breast cancer; therapy;  
KW oestrogen receptor-negative breast cancer; lung cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO9958675-A2.  
XX  
PD 18-NOV-1999.  
XX  
PF 13-MAY-1999; 99WO-US10602.  
XX  
PR 14-MAY-1998; 98US-0085426.  
PR 15-MAY-1998; 98US-0085537.

PR 15-MAY-1998; 98US-0085696.  
PR 21-OCT-1998; 98US-0105234.  
PR 27-OCT-1998; 98US-0105877.  
XX  
XX (CHIR) CHIRON CORP.  
PA (HYSE-) HYSEQ INC.  
XX  
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;  
PI Lamson G, Drmanac R, Chtvenjakov R, Dickson M, Drmanac S, Labat I;  
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Steache-Crain B;  
XX  
XX WPI: 2000-126369/11.  
DR  
XX  
XX Polynucleotide library used to determine cancerous states of mammalian  
PT cells -  
PT  
PS Claim 1; Page 445; 1097pp; English.  
XX  
XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA  
CC libraries constructed from human colon cancer cell lines. The present  
CC invention also describes a method of detecting differentially expressed  
CC genes correlated with a cancerous state of a mammalian cell, comprising  
CC detecting at least one differentially expressed gene product in a test  
CC sample derived from a cell suspected of being cancerous, where detection  
CC of the differentially expressed gene product is correlated with a  
CC cancerous state of the cell from which the test sample was derived.  
CC The polynucleotide sequences can be used in a method for detecting  
CC differentially expressed genes correlated with a cancerous state of a  
CC mammalian cell. The polynucleotides can also be used as probes for  
CC detecting and mapping related genes. They can be used in diagnosis and  
CC prognosis of diseases and disorders (e.g. identification of  
CC pre-metastatic or metastatic cancerous states, stages of cancer, or  
CC responsiveness of cancer to therapy). This is particularly for breast  
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-  
CC negative breast cancer, lung cancer, and colon cancer.  
XX  
SQ Sequence 300 BP; 128 A; 44 C; 50 G; 78 T; 0 other;

Query Match 2.5%; Score 21; DB 21; Length 300;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 215 ATGAAGTGATCTCATTTGA 235  
DB 196 ATGAAGTGATCTCATTTGA 176

RESULT 5  
ABX20395  
ID ABX20395 standard; cDNA; 390 BP.  
XX  
AC ABX20395;  
XX  
DT 10-FEB-2003 (first entry)  
XX  
DE Human GDP-mannose 4,6-dehydratase (GM4, 6D) DNA #2452.  
XX  
XX Human; GDP-mannose 4,6-dehydratase; GM4, 6D; gene; ss; inflammation;  
KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;  
KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;  
KW complex carbohydrate; gene replacement therapy; immunosuppressive;  
KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;  
KW antiaesthetic; vasotropic.  
XX  
OS Homo sapiens.  
XX  
PN US2002110548-A1.  
XX  
PD 15-AUG-2002.  
XX  
PF 11-JUN-2001; 2001US-0878574.  
XX

PR 22-NOV-1996; 96US-0753233.  
PR 03-DEC-1997; 97US-0984246.  
PR 09-SEP-1998; 98US-0149674.  
PR 14-JUN-1999; 99US-0333177.  
XX  
PA (GEM) GENETICS INST INC.  
XX  
PI Sullivan F, Kriz R, Kumar R;  
XX  
DR WPI; 2003-066673/06.  
XX  
PT New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)  
PT peptide, for manufacturing complex carbohydrates, or as targets for  
PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant  
PT rejection -  
XX  
PS Disclosure; SEQ ID NO 2454; 6pp; English.  
XX  
XX The invention relates to a composition comprising a human GDP-mannose  
CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying  
CC 4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation  
CC in a mammalian subject and for treating or ameliorating diseases affected  
CC by the level of cellular fucosylation or diseases affected by the  
CC fucosylation of glycoconjugates. These diseases include arthritis,  
CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or  
CC infection. The GM4,6D peptide or a polynucleotide encoding it is also  
CC useful for manufacturing complex carbohydrates and as targets for  
CC screening small molecule antagonists of the activity of the enzyme. The  
CC polynucleotide is useful in developing an assay for defects in the  
CC enzyme, as well as in gene replacement therapy. Sequences  
CC ABX17942-ABX17944 and ABX17947-ABX3716 represent DNA molecules encoding  
CC human GM4,6D peptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 390 BP; 141 A; 59 C; 101 G; 88 T; 1 other;  
XX  
Query Match 2.5%; Score 21; DB 25; Length 390;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 304 GATGTGAAAAAATATTGGAAC 324  
DB 300 GATGTGAAAAAATATTGGAAC 320  
XX  
RESULT 6  
AAA02459/c  
ID AAA02459 standard; cDNA; 765 BP.  
XX  
AC AAA02459;  
XX  
DT 19-MAY-2000 (first entry)  
XX  
DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:2450.  
XX  
KW Human; colon cancer; tumour; diagnosis; gene expression product;  
KW probe; detection; cancerous state; metastasis; identification;  
KW breast cancer; oestrogen receptor-positive breast cancer; therapy;  
KW oestrogen receptor-negative breast cancer; lung cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO958675-A2.  
XX  
PD 18-NOV-1999.  
XX  
PF 13-MAY-1999; 99WO-US10602.  
XX  
PR 14-MAY-1998; 98US-0085426.  
PR 15-MAY-1998; 98US-0085537.  
PR 15-MAY-1998; 98US-0085696.  
XX

PR 21-OCT-1998; 98US-0105234.  
PR 27-OCT-1998; 98US-0105877.  
XX  
PA (CHIR) CHIRON CORP.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Williams LT, Escobedo J, Imms MA, Garcia PD, Sudduth-Klinger J;  
PI Reinhard C, Glese K, Randazzo F, Kennedy GC, Pot D, Kassam A;  
PI Lamson G, Drmanac R, Ckvenjakov R, Dickson M, Drmanac S, Labat I;  
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;  
XX  
DR WPI; 2000-126369/11.  
XX  
PT Polynucleotide library used to determine cancerous states of mammalian  
PT cells -  
XX  
PS Claim 1; Page 982; 1097pp; English.  
XX  
XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA  
CC libraries constructed from human colon cancer cell lines. The present  
CC invention also describes a method of detecting differentially expressed  
CC genes correlated with a cancerous state of a mammalian cell, comprising  
CC detecting at least one differentially expressed gene product in a test  
CC sample derived from a cell suspected of being cancerous, where detection  
CC of the differentially expressed gene product is correlated with a  
CC cancerous state of the cell from which the test sample was derived.  
CC The polynucleotide sequences can be used in a method for detecting  
CC differentially expressed genes correlated with a cancerous state of a  
CC mammalian cell. The polynucleotides can also be used as probes for  
CC detecting and mapping related genes. They can be used in diagnosis and  
CC prognosis of diseases and disorders (e.g. identification of  
CC pre-metastatic or metastatic cancerous states, stages of cancer, or  
CC responsiveness of cancer to therapy). This is particularly for breast  
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-  
CC negative breast cancer, lung cancer, and colon cancer.  
XX  
SQ Sequence 765 BP; 298 A; 123 C; 136 G; 199 T; 9 other;  
XX  
Query Match 2.5%; Score 21; DB 21; Length 765;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 215 ATGAAGTGATCTCATTTCGA 235  
DB 210 ATGAAGTGATCTCATTTCGA 190  
XX  
RESULT 7  
AAT42063/c  
ID AAT42063 standard; DNA; 1830121 BP.  
XX  
AC AAT42063;  
XX  
DT 14-SEP-1999 (first entry)  
XX  
DE Haemophilus influenzae complete genome sequence.  
XX  
KW Genome; bacterium; Haemophilus influenzae; computer readable medium;  
KW expression modulating fragment; regulation; gene expression; vector;  
KW organism; open reading frame; ORF; ds.  
XX  
OS Haemophilus influenzae.  
XX  
PN WO9633276-A1.  
XX  
PD 24-OCT-1996.  
XX  
PF 22-APR-1996; 96WO-US05320.  
XX  
PR 07-JUN-1995; 95US-0487429.  
PR 21-APR-1995; 95US-0426787.  
PR 07-JUN-1995; 95US-0476102.  
XX



PA (HUMA-) HUMAN GENOME SCI INC.  
PA (UYJO ) UNIV JOHNS HOPKINS.  
PI Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;  
XX WPI; 1996-485782/48.  
XX  
XX Hemophilus influenzae Rd genome recorded on computer readable  
PT medium - useful for identifying commercially important nucleic acid  
XX fragments by homology searching  
XX  
PS Claim 1; Page 77.2-77.1091; 1291bp; English.  
XX  
CC This sequence represents the complete genome sequence of the bacterium  
CC Haemophilus influenzae strain Rd. The invention relates to a computer  
CC readable medium (CRM) having recorded upon it the complete H. influenzae  
CC nucleotide sequence (1), a representative fragment of (1) or a nucleotide  
CC sequence at least 9% identical to (1). By providing the full-length  
CC genomic sequence in a computer readable form, it is possible to identify  
CC commercially important nucleic acid fragments and expression modulating  
CC fragments (EMFs) of the Haemophilus genome. The EMFs can be used to  
CC regulate the expression of a nucleic acid molecule. Vectors and altered  
CC organisms comprising the predicted ORFs can be used to produce any of the  
CC polypeptide fragments of the H. influenzae Rd genome.  
XX  
SQ Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;  
XX  
Query Match 2.5%; Score 21; DB 17; Length 1830121;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 788 TTCTGCGAATTGACTTAT 808  
Db 1174328 TTCTGCGAATTGACTTAT 1174308  
XX  
RESULT 8  
AB212316  
ID AB212316 standard; DNA; 741 BP.  
XX  
AC AB212316;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 121.  
XX  
KM Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200216655-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 24-AUG-2001; 2001WO-US26685.  
XX  
PR 24-AUG-2000; 2000US-227866P.  
PR 26-JAN-2001; 2001US-264647P.  
PR 22-JUN-2001; 2001US-300111P.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Harper JF, Krepe J, Wang X, Zhu T;  
XX  
DR WPI; 2002-304127/34.  
XX  
PT Identifying a stress condition to which a plant cell has been exposed  
PT and producing plants with increased tolerance to these abiotic stresses  
XX  
PS Claim 144; SEQ ID NO 121; 577bp + Sequence Listing; English.  
XX

CC The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising:  
CC (a) contacting nucleic acid representative of expressed polynucleotides  
CC in the plant cell with an array or probes representative of the plant  
CC cell genome; and  
CC (b) detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used  
CC in methods of the invention.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification but is based on sequence information supplied to Derwent by  
CC the European Patent Office.  
XX  
SQ Sequence 741 BP; 251 A; 141 C; 168 G; 181 T; 0 other;  
XX  
Query Match 2.4%; Score 20; DB 24; Length 741;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 247 CTCTAGCGACAGATGCTC 266  
Db 250 CTCTAGCGACAGATGCTC 269  
XX  
RESULT 9  
AAC41093  
ID AAC41093 standard; DNA; 859 BP.  
XX  
AC AAC41093;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 30621.  
XX  
KM Hybridisation assay; genetic mapping; gene expression control;  
KM protein identification; signal transduction pathway;  
XX  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 26-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.



PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0143768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0136029.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 14-JUN-1999; 99US-0138847.  
 PR 16-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 17-JUN-1999; 99US-0139453.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 26-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0143542.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 15-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149910.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 22-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157665.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159285.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match 2.4%; Score 20; DB 21; Length 859;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 247 CTCTTAGGCAACAGATGCTC 266  
DB 359 CTCTTAGGCAACAGATGCTC 378  
  
RESULT 10  
ID AB067196 standard; DNA; 684707 BP.  
AC AB067196;  
XX 29-AUG-2002 (first entry)  
DT 29-AUG-2002 (first entry)  
XX  
DE Listeria innocua contig DNA sequence #9.  
XX  
KW Antibacterial; Listeria; food contamination; mutational analysis;  
infection; ds.  
XX  
OS Listeria innocua.  
XX  
PN WO200228891-A2.  
XX 11-APR-2002.  
PD 11-APR-2002.  
XX  
PE 04-OCT-2001; 2001WO-FR03061.  
PF 04-OCT-2001; 2001WO-FR03061.  
XX  
PR 04-OCT-2000; 2000FR-0012697.  
PS  
XX  
PA (INSP ) INST PASTEUR.  
XX (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
PI Kunst F, Glaser P;  
XX  
DR WPI; 2002-332479/37.  
XX  
XX New genomic sequences from Listeria species, useful for detection,  
PT treatment and prevention of infection, also related polypeptides,  
PT antibodies and modulators -  
XX  
PS Claim 5; SEQ ID 9; 180pp; French.  
XX  
XX The present invention relates to nucleic acid sequences  
CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes  
CC and primers for identification and/or detection of Listeria (e.g. as  
CC contaminants in foods, or mutational analysis) and for analysis of  
CC gene expression. Proteins encoded by the nucleic acid sequences can be  
CC used to screen for compounds that modulate gene expression, replication  
CC and pathogenicity of Listeria (potential therapeutic agents), also for  
CC treating infections by Listeria, and are useful as immunogens in  
CC anti-Listeria vaccines.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SO Sequence 684707 BP; 213423 A; 126375 C; 133438 G; 211468 T; 3 other;

OY 327 TAACTCTTAAGAGATTAA 346  
DB 206184 TAACTCTTAAGAGATTAA 206203  
  
RESULT 11  
ID AB069245 standard; DNA; 3011208 BP.  
AC AB069245;  
XX  
XX 29-AUG-2002 (first entry)  
DT 29-AUG-2002 (first entry)  
XX  
DE Listeria innocua DNA sequence #684.  
XX  
KW Antibacterial; Listeria; food contamination; mutational analysis;  
infection; ds.  
XX  
OS Listeria innocua.  
XX  
PN WO200228891-A2.  
XX 11-APR-2002.  
PD 11-APR-2002.  
XX  
PE 04-OCT-2001; 2001WO-FR03061.  
PF 04-OCT-2001; 2001WO-FR03061.  
XX  
PR 04-OCT-2000; 2000FR-0012697.  
PS  
XX  
PA (INSP ) INST PASTEUR.  
XX (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
PI Kunst F, Glaser P;  
XX  
DR WPI; 2002-332479/37.  
XX  
XX New genomic sequences from Listeria species, useful for detection,  
PT treatment and prevention of infection, also related polypeptides,  
PT antibodies and modulators -  
XX  
PS Claim 5; SEQ ID 2058; 180pp; French.  
XX  
XX The present invention relates to nucleic acid sequences  
CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes  
CC and primers for identification and/or detection of Listeria (e.g. as  
CC contaminants in foods, or mutational analysis) and for analysis of  
CC gene expression. Proteins encoded by the nucleic acid sequences can be  
CC used to screen for compounds that modulate gene expression, replication  
CC and pathogenicity of Listeria (potential therapeutic agents), also for  
CC treating infections by Listeria, and are useful as immunogens in  
CC anti-Listeria vaccines.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SO Sequence 3011208 BP; 941651 A; 568176 C; 559189 G; 942192 T; 0 other;

Query Match 2.4%; Score 20; DB 24; Length 3011208;  
Best Local Similarity 100.0%; Pred. No. 9.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 327 TAACTCTTAAGAGATTAA 346  
DB 63422 TAACTCTTAAGAGATTAA 63403  
  
RESULT 12  
ID AAK62944 standard; CDNA; 464 BP.  
AC AAK62944;  
XX  
XX 06-NOV-2001 (first entry)  
DT

XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:8004.  
DE  
XX  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ss.  
XX  
OS Homo sapiens.  
PN WO200157182-A2.  
XX  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184654.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0215135.  
PR 30-JUN-2000; 2000US-0214886.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 05-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 14-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234224.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 02-OCT-2000; 2000US-0237041.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246529.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249247.  
PR 17-NOV-2000; 2000US-0249248.  
PR 17-NOV-2000; 2000US-0249249.  
PR 17-NOV-2000; 2000US-0249250.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 06-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.

```

PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX P-PSDB; AAM90163.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Claim 1; SEQ ID NO 8004; 3071bp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I) by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 464 BP; 155 A; 103 C; 112 G; 90 T; 4 other;
XX
XX Query Match 2.3%; Score 19; DB 22; Length 464;
XX Best Local Similarity 100.0%; Pred. No. 33;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 42 TTCATGACCTGATGAGAA 60
DB 322 TTCATGACCTGATGAGAA 304

RESULT 13
AAS22921
ID AAS22921 standard; DNA; 737 BP.
XX
XX AAS22921;
XX
XX 24-OCT-2001 (first entry)
XX
XX DNA encoding novel bone marrow polypeptide #15.
XX
XX Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
XX haematopoietic; myeloid; lymph cell disorder; tissue regeneration;
XX wound healing; nutritional supplement; immune disorder;
XX severe combined immunodeficiency; SCID; ds.
XX
XX Homo sapiens.
XX
XX MO200157187-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US03782.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 20-JUN-2000; 2000US-0598075.
XX 19-JUL-2000; 2000US-0620325.

```

```

PR 30-NOV-2000; 2000US-0250683.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ,
XX Ren F, Dymnac RT;
XX
XX WPI; 2001-488875/53.
XX
XX P-PSDB; AAU14616.
XX
XX Nucleic acids encoding bone marrow polypeptides, useful in diagnostic
XX and gene therapy -
XX
XX Claim 1; Page 162; 392bp; English.
XX
XX AAS22907-AAS23099 represent nucleic acids encoding novel bone marrow
XX polypeptides. The nucleic acids and corresponding proteins may be used
XX in the prevention, diagnosis and treatment of diseases associated with
XX inappropriate bone marrow polypeptide expression. For example, to treat
XX disorders associated with decreased expression by rectifying mutations
XX or deletions in a patient's genome that affect the activity of the
XX polypeptides by expressing inactive proteins or to supplement the
XX patient's own production of the polypeptide. Additionally, the nucleic
XX acids may be used to produce the polypeptides, by inserting the nucleic
XX acids into a host cell and culturing the cell to express the protein.
XX The nucleic acid and its complementary sequences may also be used as DNA
XX probes in diagnostic assays to detect and quantitate the presence of
XX similar nucleic acid sequences in samples, and therefore which patients
XX may be in need of restorative therapy. The proteins may also be used as
XX antigens in the production of antibodies against bone marrow proteins
XX and in assays to identify modulators of their expression and activity.
XX The anti-bone marrow protein antibodies and antagonists may also be used
XX to down regulate expression and activity. The antibodies may also be used
XX as diagnostic agents for detecting the presence of the protein in samples
XX (e.g. by enzyme linked immunosorbent assay (ELISA)). The proteins
XX may be used to regulate haematopoietic activity, and consequently in the
XX treatment of myeloid or lymph cell disorders; in tissue regeneration,
XX such as wound healing; as a nutritional supplement; and in treatment of
XX immune disorders such as severe combined immunodeficiency (SCID).
XX
XX Sequence 737 BP; 232 A; 112 C; 182 G; 191 T; 20 other;
XX
XX Query Match 2.3%; Score 19; DB 22; Length 737;
XX Best Local Similarity 100.0%; Pred. No. 33;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 617 TTGTAATTATTTTGAAA 635
DB 591 TTGTAATTATTTTGAAA 609

RESULT 14
AA70171/c
ID AA70171 standard; DNA; 810 BP.
XX
XX AA70171;
XX
XX 07-NOV-2000 (first entry)
XX
XX Plasmidium falciparum chromosome 2 related DNA sequence SEQ ID NO:304.
XX
XX Plasmidium falciparum; chromosome 2; human malaria parasite; vaccine;
XX antimalarial; malarial; protozoacide; infection; insecticide; ds.
XX
XX Plasmidium falciparum.
XX
XX MO200025728-A2.
XX
XX 11-MAY-2000.
XX
XX 05-NOV-1999; 99WO-US26796.
XX 05-NOV-1998; 98US-0107131.

```

XX (HOFF/) HOFFMAN S.  
 PA (CARU/) CARUCCI D.  
 PA (GARD/) GARDNER M.  
 PA (VENT/) VENTER J C.  
 XX  
 PI Hoffman S, Carucci D, Gardner M, Venter JC;  
 DR WPI: 2000-365347/31.  
 XX  
 PT Proteins encoded by chromosome 2 of the human malarial parasite,  
 PT Plasmodium falciparum, useful as antimalarial vaccines and in the  
 PT diagnosis of P.falciparum infection -  
 PS Disclosure; Page 508; 577pp; English.  
 XX  
 CC The present invention describes proteins and their fragments (I) encoded  
 CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.  
 CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)  
 CC vaccines against P. falciparum infection comprising (I) or (II).  
 CC (I) and (II) are useful for the development of vaccines against  
 CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal  
 CC antibody raised to immunogens comprising the sequences of (I), are  
 CC useful in the detection of infection with P. falciparum. Furthermore,  
 CC (I) (especially when they are rifins or secreted or membrane proteins)  
 CC can aid the identification of drugs to treat or prevent P. falciparum  
 CC infection, or they can be used to identify drug resistance in  
 CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the  
 CC subsequent identification of proteins encoded by it will help to expand  
 CC our understanding of parasite biology, a process hampered by the  
 CC complexity of the parasitic life cycle, and provide new targets for  
 CC vaccine and drug development. Parasite resistance to drugs and mosquito  
 CC resistance to insecticides have led to a resurgence of malaria in many  
 CC parts of the world, and there is a pressing need for vaccines and new  
 CC drugs. AAH70078 to AAH70287 and AAB18144 to AAB18352 represent nucleotide  
 CC and protein sequences given in the present invention, but which are not  
 CC specifically mentioned within the specification.  
 XX  
 SQ Sequence 810 BP; 325 A; 81 C; 96 G; 308 T; 0 other;  
 Query Match 2.3%; Score 19; DB 21; Length 810;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 497 AACATTCGACGAAATAT 515  
 DB 153 AACATTCGACGAAATAT 135  
 RESULT 15  
 AAD05784  
 ID AAD05784 standard; cDNA; 1137 BP.  
 XX  
 AC AAD05784;  
 XX  
 DT 31-JUL-2001 (first entry)  
 XX  
 DE Arabidopsis thaliana transcription factor, G1324 cDNA.  
 XX  
 KW Transcription factor; biochemical characteristic; controlling element;  
 KW structural characteristic; developmental characteristic; gene therapy;  
 KW agricultural biotechnology; plant trait modification; ss.  
 OS Arabidopsis thaliana.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 54..914  
 FT /tag=a  
 FT /product="Transcription factor, G1324"  
 XX  
 PN MO200136597-A1.  
 XX  
 PD 25-MAY-2001.

XX 14-NOV-2000; 2000MO-US31344.  
 XX  
 XX 17-NOV-1999; 99US-0166228.  
 PR 17-APR-2000; 2000US-0197699.  
 PR 22-AUG-2000; 2000US-0227439.  
 XX  
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (CREE/) CREELMAN R.  
 PA (YUGG/) YU G.  
 PA (ADAM/) ADAM L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (HEAR/) HEARD J.  
 PA (SAMA/) SAMAHA R.  
 PA (PLIG/) PILGRIM M.  
 PA (PINE/) PINEDA O.  
 PA (JIANG/) JIANG C.  
 XX  
 PI Creelman R, Yu G, Adam L, Riechmann JL, Heard J, Samaha R;  
 PI Pilgrim M, Pineda O, Jiang C;  
 DR WPI: 2001-335999/35.  
 DR P-PSDB; AAB01900.  
 XX  
 PT Nucleic acids encoding plant transcription factor polypeptides, useful  
 PT for altering the biochemical characteristics of plants e.g. corn,  
 PT potato and cotton plants -  
 PS Claim 4; Page 87-88; 127pp; English.  
 XX  
 CC The present sequence is Arabidopsis thaliana transcription factor,  
 CC G1324 cDNA. The transcription factor is used for altering a plant's  
 CC biochemical characteristics. The transcription factor may be used to  
 CC alter the structure and developmental characteristics of plants such as  
 CC soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,  
 CC alfalfa, sugar cane, turf, banana, blackberry, blueberry, strawberry,  
 CC raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant,  
 CC grapes, honey dew, lettuce, mango, melon, onion, papaya, peas, peppers,  
 CC pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon,  
 CC roseaceous fruits and/or vegetable brassicas. Transcription factors are  
 CC key controlling elements of biological pathways and altering expression  
 CC levels of 1 or more transcription factors can change entire biological  
 CC pathways in an organism. Therefore manipulating transcription factor  
 CC levels in plants offers great potential in agricultural biotechnology  
 CC for modifying a plant's traits. Transcription factor cDNA is useful in  
 CC gene therapy.  
 XX  
 SQ Sequence 1137 BP; 378 A; 213 C; 233 G; 313 T; 0 other;  
 Query Match 2.3%; Score 19; DB 22; Length 1137;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 146 GGAAGTGTAGATTGAG 164  
 DB 223 GGAAGTGTAGATTGAG 241

Search completed: January 30, 2004, 09:36:34  
 Job time : 282 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 09:23:58 ; Search time 1953 Seconds  
(without alignments)  
10266.868 Million cell updates/sec

Title: US-10-033-190-1

Perfect score: 825

Sequence: 1 atgaacagacatcatatgc.....tatgaatctactatgtaa 825

Scoring table: OLIGO\_NUC

Searched: 22781392 seqs, 1215238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_rtd:\*  
27: em\_gss\_vrt:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4.2	371	10	BF646896	BF646896 NF077D01E
2	4.2	671	10	BG447824	BG447824 NF103E12E
3	4.1	335	10	BE462229	BE462229 EST324493
4	4.1	418	9	AM626121	AM626121 EST320028

5	4.1	490	10	BE462282	BE462282 EST324546
6	3.4	245	10	BF519256	BF519256 EST456718
7	3.3	367	10	BF635572	BF635572 NF104H01D
8	3.3	526	10	AM980341	AM980341 EST391494
9	3.3	526	10	BF006562	BF006562 EST435060
10	3.3	549	10	BE124179	BE124179 EST394304
11	3.3	585	10	BG588067	BG588067 EST4509558
12	3.3	618	10	BG647939	BG647939 EST509558
13	3.3	651	10	BG647226	BG647226 EST508845
14	3.2	781	13	BH830456	BH830456 T008R08 P
15	3.0	308	13	BH876538	BH876538 V021H11 P
16	2.9	660	14	CD038321	CD038321 UTPPT005
17	2.8	510	10	BF325282	BF325282 su20e03.y
18	2.7	154	28	BH008948	BH008948 e122a03.x
19	2.5	494	14	CA410578	CA410578 843 F -P
20	2.5	531	13	BH926469	BH926469 aa875c05.
21	2.5	588	29	BZ345026	BZ345026 ht43g03.b
22	2.5	595	12	BM062915	BM062915 KS01048H0
23	2.5	608	29	AG242336	AG242336 Lotur j4p
24	2.5	630	12	BM067328	BM067328 KS08003A0
25	2.5	635	12	BI095605	BI095605 ACS27-JM1
26	2.5	636	13	BH927658	BH927658 K006P59P
27	2.5	655	9	AM616647	AM616647 EST323058
28	2.5	677	12	BM063980	BM063980 KS01061P1
29	2.5	696	12	BM064030	BM064030 KS01062C0
30	2.5	767	10	BG435775	BG435775 602506882
31	2.5	900	29	CNS07ACM	AL436412 T7 end of
32	2.4	425	14	CD480621	CD480621 eca01-18m
33	2.4	426	29	BZ705431	BZ705431 PUBMT69TD
34	2.4	460	13	BQ848932	BQ848932 QGAB107.y
35	2.4	481	13	BU998080	BU998080 PESTroab9
36	2.4	489	12	BM293101	BM293101 EST575643
37	2.4	557	13	BM179473	BM179473 BM179473
38	2.4	558	28	AQ942653	AQ942653 Sheared.D
39	2.4	565	14	CA589420	CA589420 hab58a03.
40	2.4	584	28	AZ415511	AZ415511 IM0190G19
41	2.4	587	12	BQ036614	BQ036614 SL5-0172
42	2.4	596	12	BM291189	BM291189 EST577723
43	2.4	597	9	AV980999	AV980999 AV980999
44	2.4	618	9	AM685354	AM685354 NF026E09N
45	2.4	636	13	BM260274	BM260274 BM260274
46	2.4	639	9	AM223368	AM223368 EST300179
47	2.4	647	9	AV997065	AV997065 AV997065
48	2.4	664	10	BG125772	BG125772 EST471418
49	2.4	672	13	BM223289	BM223289 BM223289
50	2.4	678	28	BH163620	BH163620 ENTS04TF
51	2.4	683	9	AV674350	AV674350 AV674350
52	2.4	700	29	BZ600408	BZ600408 WHABD82TF
53	2.4	715	9	AV976001	AV976001 AV976001
54	2.4	756	14	CA342196	CA342196 672018 NC
55	2.4	774	14	BY756830	BY756830 BY756830
56	2.4	782	28	BZ080728	BZ080728 11c05C09.
57	2.4	782	28	BZ080824	BZ080824 11c05C09.
58	2.4	783	13	BM074536	BM074536 BM074536
59	2.4	784	13	BM093209	BM093209 BM093209
60	2.4	796	13	BM063421	BM063421 BM063421
61	2.4	835	29	BZ428248	BZ428248 BONTV52TR
62	2.4	1381	10	BG563853	BG563853 602584657
63	2.3	225	9	AV418859	AV418859 AV418859
64	2.3	229	13	BO600886	BO600886 MI-P-E7-a
65	2.3	247	28	AZ262504	AZ262504 RPCI-23-1
66	2.3	261	12	BJ082401	BJ082401 BJ082401
67	2.3	330	10	BE491490	BE491490 db71f06.x
68	2.3	333	13	BU743442	BU743442 mai31h08.
69	2.3	346	14	W19124	W19124 zb15c02.tl
70	2.3	348	9	AJ499746	AJ499746 AJ499746
71	2.3	358	9	AV422244	AV422244 AV422244
72	2.3	359	9	AV422287	AV422287 AV422287
73	2.3	360	9	AM203878	AM203878 SWAMCA46
74	2.3	366	13	BQ740451	BQ740451 baq47e04.
75	2.3	394	12	BJ048318	BJ048318 BJ048318
76	2.3	396	9	AA206774	AA206774 zq80c01.s
77	2.3	407	9	AM636175	AM636175 b143fE10.w

78	19	2.3	417	28	BZ145123	BZ145123	CH230-319	151	19	2.3	694	28	BH952322	BH952322	cd185e04.
79	19	2.3	421	9	AV419261	AV419261		152	19	2.3	700	12	BM168641	BM168641	EST571164
80	19	2.3	442	10	BE521395	BE521395	M19D1STM	153	19	2.3	704	13	BM206615	BM206615	EST206615
81	19	2.3	460	9	A1757853	A1757853	EST623148	154	19	2.3	718	12	BM170240	BM170240	EST572763
82	19	2.3	468	13	BQ515733	BQ515733	EST623148	155	19	2.3	728	10	BP232274	BP232274	de01e06.x
83	19	2.3	472	10	BF596448	BF596448	BU51405.Y	156	19	2.3	731	12	BM164453	BM164453	EST566965
84	19	2.3	475	28	AZ236258	AZ236258	RPCI-23-8	157	19	2.3	732	12	BM170084	BM170084	EST572607
85	19	2.3	476	10	BE471915	BE471915	EST416768	158	19	2.3	735	12	BM166185	BM166185	EST568708
86	19	2.3	478	29	BZ274896	BZ274896	CH230-307	159	19	2.3	744	28	AO868872	AO868872	nbe0033K
87	19	2.3	480	9	A1455246	A1455246	LD2252.3	160	19	2.3	747	13	BO514458	BO514458	EST621873
88	19	2.3	484	29	BZ515514	BZ515514	BDMR157TR	161	19	2.3	747	28	BM126277	BM126277	BARC-Sat
89	19	2.3	489	13	BQ515732	BQ515732	EST623147	162	19	2.3	753	28	BM166330	BM166330	BOH8221F
90	19	2.3	491	12	BI433409	BI433409	EST536170	163	19	2.3	772	12	BM166640	BM166640	EST569163
91	19	2.3	496	9	AM617393	AM617393	EST323804	164	19	2.3	813	28	AO746554	AO746554	HS-2278.A
92	19	2.3	500	28	AZ906608	AZ906608	RPCI-24-1	165	19	2.3	821	29	CC084429	CC084429	CSU-K33F.
93	19	2.3	511	10	BF453355	BF453355	nab38e05.	166	19	2.3	836	10	BF668566	BF668566	602123613
94	19	2.3	514	12	BI417595	BI417595	LJNEST33D	167	19	2.3	841	10	BE877239	BE877239	601485163
95	19	2.3	516	12	BM165411	BM165411	EST567934	168	19	2.3	845	29	CC079179	CC079179	CSU-K33F.
96	19	2.3	516	14	CB076834	CB076834	h145C07.G	169	19	2.3	847	29	BZ432618	BZ432618	BONR221F
97	19	2.3	522	9	AM102437	AM102437	bd88a10.Y	170	19	2.3	944	13	CA983350	CA983350	AGENCOURT
98	19	2.3	522	10	BF060655	BF060655	7159P10.X	171	19	2.3	945	14	BO926786	BO926786	AGENCOURT
99	19	2.3	524	13	BQ099899	BQ099899	dp69H08.Y	172	19	2.3	975	10	BG355237	BG355237	602369801
100	19	2.3	524	14	CB941372	CB941372	AGENCOURT	173	19	2.3	1238	13	BU235640	BU235640	603412115
101	19	2.3	527	12	BI359086	BI359086	949034A11	174	19	2.3	2862	11	AK034745	AK034745	Mus muscu
102	19	2.3	533	13	BQ833086	BQ833086	L61n1082	175	19	2.3	3113	11	AK084655	AK084655	Mus muscu
103	19	2.3	534	9	AV604131	AV604131	AV604131	176	19	2.3	3205	11	AK046812	AK046812	Mus muscu
104	19	2.3	534	9	AM220176	AM220176	EST302659	177	18	2.2	106	10	BG206618	BG206618	RST26069
105	19	2.3	539	10	BG227110	BG227110	hp97P10.Y	178	18	2.2	153	28	AO034816	AO034816	CIT-HSP-2
106	19	2.3	542	28	AO210636	AO210636	HS-2225.B	179	18	2.2	166	28	BH337605	BH337605	701553307
107	19	2.3	545	13	BQ830215	BQ830215	L61n12133	180	18	2.2	200	9	AI999296	AI999296	701553307
108	19	2.3	546	9	AM035257	AM035257	EST280519	181	18	2.2	206	10	BG193266	BG193266	RST12393
109	19	2.3	556	10	AM942910	AM942910	LD28817.3	182	18	2.2	206	10	BG202400	BG202400	RST12393
110	19	2.3	557	28	AZ095011	AZ095011	RPCI-23-4	183	18	2.2	207	14	CB469987	CB469987	sn14.E12.
111	19	2.3	566	12	BM145657	BM145657	EST567090	184	18	2.2	208	10	BG186977	BG186977	RST5955.A
112	19	2.3	567	12	BM160260	BM160260	EST562783	185	18	2.2	212	14	Z48422	Z48422	ATTS4636.G1
113	19	2.3	571	10	AV957319	AV957319	AV957319	186	18	2.2	213	9	N28000	N28000	YK49H03.r1
114	19	2.3	573	10	BG022036	BG022036	df95G11.X	187	18	2.2	228	10	BG627273	BG627273	CC-estf1CL
115	19	2.3	578	9	AV977051	AV977051	AV977051	188	18	2.2	230	29	CC348828	CC348828	OGIAK89TH
116	19	2.3	583	9	A1498826	A1498826	AV498826	189	18	2.2	231	10	BG627272	BG627272	CC-estf1CL
117	19	2.3	583	10	BG130989	BG130989	EST465881	190	18	2.2	238	12	BU122924	BU122924	B122924
118	19	2.3	583	10	BE187553	BE187553	EST336114	191	18	2.2	239	12	BM160482	BM160482	EST5633005
119	19	2.3	586	14	CA524237	CA524237	K512034F0	192	18	2.2	240	9	AV203809	AV203809	AZ0203809
120	19	2.3	591	13	BU966512	BU966512	rF04G04.Y	193	18	2.2	242	9	AU201570	AU201570	AU201570
121	19	2.3	592	14	CA809843	CA809843	CA221011	194	18	2.2	245	9	AV414695	AV414695	AV414695
122	19	2.3	595	13	BU779023	BU779023	SJESER10	195	18	2.2	245	12	BM156758	BM156758	fV558602.X
123	19	2.3	596	12	BI449376	BI449376	de627E01.X	196	18	2.2	246	9	AV957931	AV957931	AV957931
124	19	2.3	600	14	CAS30577	CAS30577	9043-72.M	197	18	2.2	252	9	AJ560011	AJ560011	AJ560011
125	19	2.3	605	28	AZ367488	AZ367488	IM0117109	198	18	2.2	252	13	BU030717	BU030717	OHU16D12.
126	19	2.3	606	9	AV957845	AV957845	AV957845	199	18	2.2	252	28	AO070422	AO070422	HS-3035.B
127	19	2.3	607	13	BQ513598	BQ513598	EST621013	200	18	2.2	261	13	BO901933	BO901933	Ta02.1060
128	19	2.3	612	13	BO125515	BO125515	rc69C09.Y	201	18	2.2	272	13	BU031227	BU031227	OHU17M24.
129	19	2.3	617	10	BE508713	BE508713	dp61C06.X	202	18	2.2	274	13	BU034866	BU034866	OHU5G10.Y
130	19	2.3	622	13	BQ386261	BQ386261	NJSC_mml7	203	18	2.2	276	13	BU032274	BU032274	OHU20118.
131	19	2.3	624	10	BE508714	BE508714	db61d05.X	204	18	2.2	276	29	BZ926045	BZ926045	CH240-73A
132	19	2.3	625	28	BI417740	BI417740	LJNEST34G	205	18	2.2	281	9	AM015330	AM015330	UI-H-B10-
133	19	2.3	625	12	AZ368260	AZ368260	IM0118F16	206	18	2.2	281	12	BM872921	BM872921	mgmY005X1
134	19	2.3	625	14	CA994863	CA994863	rc20a03.Y	207	18	2.2	281	9	AA815564	AA815564	vb97C02.X
135	19	2.3	627	10	BG126114	BG126114	EST471760	208	18	2.2	281	12	BM872921	BM872921	mgmY005X1
136	19	2.3	632	9	AV963611	AV963611	AV963611	209	18	2.2	281	28	AO666756	AO666756	HS-2104.B
137	19	2.3	641	29	CC094742	CC094742	CSU-K34.1	210	18	2.2	281	10	BG189076	BG189076	RST8112.A
138	19	2.3	650	28	B2127034	B2127034	CH230-309	211	18	2.2	295	13	BO834548	BO834548	ALEST0512
139	19	2.3	658	12	BM159766	BM159766	EST562289	212	18	2.2	300	13	BO976117	BO976117	OHU120112.
140	19	2.3	661	13	BM263176	BM263176	BM263176	213	18	2.2	300	14	C07272	C07272	C07272
141	19	2.3	662	13	BM233321	BM233321	BM233321	214	18	2.2	304	9	AA815564	AA815564	zab9f09.X
142	19	2.3	665	12	BI1308864	BI1308864	EST530274	215	18	2.2	316	28	AO923627	AO923627	RPCI-23-2
143	19	2.3	673	12	BM163307	BM163307	EST565830	216	18	2.2	324	28	AZ012415	AZ012415	RPCI-23-3
144	19	2.3	674	28	BH477222	BH477222	BOGML34TF	217	18	2.2	330	14	H82391	H82391	yV79G08.r1
145	19	2.3	674	10	BI082116	BI082116	BI082116	218	18	2.2	333	9	AM031724	AM031724	EST275178
146	19	2.3	678	10	BB619615	BB619615	BB619615	219	18	2.2	339	9	AV428856	AV428856	AV428856
147	19	2.3	683	13	BM260265	BM260265	BM260265	220	18	2.2	339	28	AZ502281	AZ502281	IM0341D24
148	19	2.3	683	28	BH523564	BH523564	BOHAT26TF	221	18	2.2	341	9	AU109735	AU109735	AMU09735
149	19	2.3	684	10	BF632342	BF632342	NF017604D	222	18	2.2	342	14	CB693607	CB693607	AMGNNUC.C
150	19	2.3	684	28	BZ028604	BZ028604	oe85H04.	223	18	2.2	342	14	CB693607	CB693607	AMGNNUC.C



c 224	18	2.2	344	9	AU110269	AU110269	c 297	18	2.2	422	9	AV423938
225	18	2.2	345	10	BB251186	BB251186	c 298	18	2.2	422	28	A2119787
226	18	2.2	346	9	AV955104	AV955104	c 299	18	2.2	423	9	AM567149
c 227	18	2.2	354	10	BG124448	BG124448	c 300	18	2.2	423	10	BE268190
c 228	18	2.2	357	10	BG124979	BG124979	c 301	18	2.2	423	14	CA410518
229	18	2.2	359	9	AM736494	AM736494	c 302	18	2.2	429	9	AI897109
c 230	18	2.2	360	9	AV189489	AV189489	c 303	18	2.2	429	9	AV792025
c 231	18	2.2	360	9	AV189774	AV189774	c 304	18	2.2	429	29	CNS00091
c 232	18	2.2	360	9	AV192178	AV192178	c 305	18	2.2	433	9	AV797384
c 233	18	2.2	360	9	AV201775	AV201775	c 306	18	2.2	434	10	BE923104
c 234	18	2.2	360	9	AV203603	AV203603	c 307	18	2.2	435	9	AI756470
c 235	18	2.2	360	14	CA3200	CA3200	c 308	18	2.2	435	14	CD474189
c 236	18	2.2	360	14	CA6124	CA6124	c 309	18	2.2	435	28	BH734914
c 237	18	2.2	360	14	CA6365	CA6365	c 310	18	2.2	436	10	BE200612
c 238	18	2.2	360	14	CA69107	CA69107	c 311	18	2.2	436	10	BF100446
c 239	18	2.2	360	14	CA6954	CA6954	c 312	18	2.2	436	28	AO436937
c 240	18	2.2	360	14	CA6986	CA6986	c 313	18	2.2	437	9	AA844518
c 241	18	2.2	362	14	CA6986	CA6986	c 314	18	2.2	437	9	AA844518
c 242	18	2.2	362	14	CA6986	CA6986	c 315	18	2.2	438	28	A2910121
c 243	18	2.2	363	9	AA223752	AA223752	c 316	18	2.2	439	10	BE207628
c 244	18	2.2	363	9	AA223752	AA223752	c 317	18	2.2	439	12	BM360403
c 245	18	2.2	364	29	AG252754	AG252754	c 318	18	2.2	439	28	AO826539
c 246	18	2.2	365	10	BG198275	BG198275	c 319	18	2.2	439	28	AO215381
c 247	18	2.2	366	14	CB386340	CB386340	c 320	18	2.2	442	10	BG198806
c 248	18	2.2	370	10	BG188016	BG188016	c 321	18	2.2	443	10	AM852631
c 249	18	2.2	370	10	BG202401	BG202401	c 322	18	2.2	444	28	AO817618
c 250	18	2.2	370	10	BG202966	BG202966	c 323	18	2.2	448	13	BM204407
c 251	18	2.2	370	10	BG213398	BG213398	c 324	18	2.2	448	28	A2004484
c 252	18	2.2	370	10	BG220205	BG220205	c 325	18	2.2	449	9	AI899497
c 253	18	2.2	371	9	AV417403	AV417403	c 326	18	2.2	449	14	CD525711
c 254	18	2.2	371	10	BG186978	BG186978	c 327	18	2.2	451	10	BG191178
c 255	18	2.2	372	28	AO613213	AO613213	c 328	18	2.2	453	10	BF600417
c 256	18	2.2	373	10	BG186479	BG186479	c 329	18	2.2	453	12	BM731482
c 257	18	2.2	373	10	BG206617	BG206617	c 330	18	2.2	453	14	R85555
c 258	18	2.2	373	10	BG218722	BG218722	c 331	18	2.2	454	14	CA410620
c 259	18	2.2	373	10	BG218722	BG218722	c 332	18	2.2	454	14	CB020777
c 260	18	2.2	374	9	AV187345	AV187345	c 333	18	2.2	456	9	AL373136
c 261	18	2.2	374	14	CA13042	CA13042	c 334	18	2.2	456	28	AU208391
c 262	18	2.2	375	14	CA6629	CA6629	c 335	18	2.2	456	28	AO913996
c 263	18	2.2	375	14	CA6629	CA6629	c 336	18	2.2	457	13	BM259430
c 264	18	2.2	376	13	BY229266	BY229266	c 337	18	2.2	459	10	BE609255
c 265	18	2.2	376	13	BY229266	BY229266	c 338	18	2.2	459	13	BQ583069
c 266	18	2.2	377	10	BE612165	BE612165	c 339	18	2.2	459	14	CD473244
c 267	18	2.2	378	14	CA69921	CA69921	c 340	18	2.2	459	28	AZ881939
c 268	18	2.2	380	9	AI210396	AI210396	c 341	18	2.2	460	9	AI781114
c 269	18	2.2	381	14	CA68581	CA68581	c 342	18	2.2	460	28	AZ885938
c 270	18	2.2	385	12	BI745542	BI745542	c 343	18	2.2	461	9	AU286822
c 271	18	2.2	388	14	CB043597	CB043597	c 344	18	2.2	463	9	AI781115
c 272	18	2.2	389	14	CB043420	CB043420	c 345	18	2.2	464	9	AI728574
c 273	18	2.2	390	9	AA214335	AA214335	c 346	18	2.2	464	9	AI731880
c 274	18	2.2	392	10	BF067735	BF067735	c 347	18	2.2	464	10	AM852656
c 275	18	2.2	395	9	AV414447	AV414447	c 348	18	2.2	465	13	BI559495
c 276	18	2.2	395	12	BG898149	BG898149	c 349	18	2.2	465	14	CB652279
c 277	18	2.2	397	28	AZ302055	AZ302055	c 350	18	2.2	467	12	BI553198
c 278	18	2.2	398	10	BE921420	BE921420	c 351	18	2.2	469	13	BUT65396
c 279	18	2.2	398	29	AI757286	AI757286	c 352	18	2.2	470	10	BE523903
c 280	18	2.2	400	14	CB043382	CB043382	c 353	18	2.2	475	9	AU221413
c 281	18	2.2	404	13	BQ380354	BQ380354	c 354	18	2.2	475	14	CD524670
c 282	18	2.2	404	13	BY070910	BY070910	c 355	18	2.2	476	14	CD524724
c 283	18	2.2	405	28	AO513493	AO513493	c 356	18	2.2	478	28	AO700855
c 284	18	2.2	406	9	AI442348	AI442348	c 357	18	2.2	479	14	CB271935
c 285	18	2.2	406	12	CA535883	CA535883	c 358	18	2.2	481	12	BM360393
c 286	18	2.2	409	12	BG883443	BG883443	c 359	18	2.2	482	9	AV955107
c 287	18	2.2	409	29	BI153243	BI153243	c 360	18	2.2	485	12	BM068672
c 288	18	2.2	410	13	BM203042	BM203042	c 361	18	2.2	486	9	AI955124
c 289	18	2.2	411	10	BF810926	BF810926	c 362	18	2.2	487	9	AI388632
c 290	18	2.2	414	12	BM315610	BM315610	c 363	18	2.2	487	9	AA446695
c 291	18	2.2	414	14	CA567876	CA567876	c 364	18	2.2	488	9	AV947540
c 292	18	2.2	417	13	BM203161	BM203161	c 365	18	2.2	488	10	BE503055
c 293	18	2.2	417	14	CA994734	CA994734	c 366	18	2.2	488	28	AZ052200
c 294	18	2.2	420	10	BF595252	BF595252	c 367	18	2.2	489	9	AI084475
c 295	18	2.2	421	10	BF595252	BF595252	c 368	18	2.2	493	11	AIY07442
c 296	18	2.2	421	12	BQ040564	BQ040564	c 369	18	2.2	494	9	AM829771
c 297	18	2.2	422	9	AV423938	AV423938	c 298	18	2.2	422	28	A2119787
c 298	18	2.2	423	9	AM567149	AM567149	c 299	18	2.2	423	10	BE268190
c 299	18	2.2	423	10	BE268190	BE268190	c 300	18	2.2	423	14	CA410518
c 300	18	2.2	423	14	CA410518	CA410518	c 301	18	2.2	429	9	AI897109
c 301	18	2.2	429	9	AV792025	AV792025	c 302	18	2.2	429	29	CNS00091
c 302	18	2.2	429	29	CNS00091	CNS00091	c 303	18	2.2	433	9	AV797384
c 303	18	2.2	433	9	AV797384	AV797384	c 304	18	2.2	434	10	BE923104
c 304	18	2.2	434	10	BE923104	BE923104	c 305	18	2.2	435	9	AI756470
c 305	18	2.2	435	9	AI756470	AI756470	c 306	18	2.2	435	14	CD474189
c 306	18	2.2	435	14	CD474189	CD474189	c 307	18	2.2	436	10	BE200612
c 307	18	2.2	436	10	BE200612	BE200612	c 308	18	2.2	436	10	BF100446
c 308	18	2.2	436	10	BF100446	BF100446	c 309	18	2.2	437	9	AA844518
c 309	18	2.2	437	9	AA844518	AA844518	c 310	18	2.2	437	9	AA844518
c 310	18	2.2	437	9	AA844518	AA844518	c 311	18	2.2	438	28	A2910121
c 311	18	2.2	438	28	A2910121	A2910121	c 312	18	2.2	439	10	BE207628
c 312	18	2.2	439	10	BE207628	BE207628	c 313	18	2.2	439	12	BM360403
c 313	18	2.2	439	12	BM360403	BM360403	c 314	18	2.2	439	28	AO826539
c 314	18	2.2	439	28	AO826539	AO826539	c 315	18	2.2	439	28	AO215381
c 315	18	2.2	439	28	AO215381	AO215381	c 316	18	2.2	442	10	BG198806
c 316	18	2.2	442	10	BG198806	BG198806	c 317	18	2.2	443	10	AM852631
c 317	18	2.2	443	10	AM852631	AM852631	c 318	18	2.2	444	28	AO817618
c 318	18	2.2	444	28	AO817618	AO817618	c 319	18	2.2	448	13	BM204407
c 319	18	2.2	448	13	BM204407	BM204407	c 320	18	2.2	448	28	A2004484
c 320	18	2.2	448	28	A2004484	A2004484	c 321	18	2.2	449	9	AI899497
c 321	18	2.2	449	9	AI899497	AI899497	c 322	18	2.2	449	14	CD525711
c 322	18	2.2	451	10	BG191178	BG191178	c 323	18	2.2	451	10	BF600417
c 323	18	2.2	453	10	BF600417	BF600417	c 324	18	2.2	453	12	BM731482
c 324	18	2.2	453	12	BM731482	BM731482	c 325	18	2.2	453	14	R85555
c 325	18	2.2	454	14	CA410620	CA410620	c 326	18	2.2	454	14	CB020777
c 326	18	2.2	454	14	CB020777	CB020777	c 327	18	2.2	456	9	AL373136
c 327	18	2.2	456	9	AL373136	AL373136	c 328	18	2.2	456	28	AU208391
c 328	18	2.2	456	28	AU208391	AU208391						

C 370	18	2.2	494	10	BG366891	BG366891	HVSMEM1000	C 443	18	2.2	557	29	DR30B12S	AL738869	Danio rer
C 371	18	2.2	495	9	AU210851	AU210851		C 444	18	2.2	558	12	BJ109738	BJ109738	BJ109738
C 372	18	2.2	495	12	BJ190056	BJ190056		C 445	18	2.2	559	9	AU286811	AU286811	AU286811
C 373	18	2.2	496	13	BQ803903	BQ803903	WHE2843.F	C 446	18	2.2	559	12	BJ108709	BJ108709	BJ108709
C 374	18	2.2	501	13	BM277507	BM277507		C 447	18	2.2	559	13	BJ779698	BJ779698	BJ779698
C 375	18	2.2	507	10	BE322300	BE322300	BF023A031	C 448	18	2.2	562	9	AM60567	AM60567	AM60567
C 376	18	2.2	508	13	AM031191	AM031191	EST274729	C 449	18	2.2	562	13	BM218743	BM218743	BM218743
C 377	18	2.2	509	9	BQ091977	BQ091977	F435d01.x	C 450	18	2.2	562	14	CB384500	CB384500	OSTF003F6
C 378	18	2.2	509	9	AU199963	AU199963		C 451	18	2.2	562	14	CD344678	CD344678	ESTE9E87
C 379	18	2.2	510	12	BJ124890	BJ124890		C 452	18	2.2	564	10	BE611947	BE611947	gr02a01.Y
C 380	18	2.2	510	28	AZ855654	AZ855654	2M0159115	C 453	18	2.2	564	10	BE922729	BE922729	EST426498
C 381	18	2.2	512	14	CA561452	CA561452	K0287C06-	C 454	18	2.2	564	10	BF194180	BF194180	
C 382	18	2.2	512	14	CA717144	CA717144	wk4C.pk0	C 455	18	2.2	564	13	BM266676	BM266676	
C 383	18	2.2	512	28	AZ882831	AZ882831	RPCT-23-1	C 456	18	2.2	566	9	AU001511	AU001511	
C 384	18	2.2	514	9	AA251625	AA251625	zsl10a09.x	C 457	18	2.2	566	9	AV960818	AV960818	
C 385	18	2.2	515	14	M43174	M43174	22569.Lambd	C 458	18	2.2	566	9	BO99707	BO99707	AV960818
C 386	18	2.2	515	12	BH000383	BH000383	2M0288M07	C 459	18	2.2	566	13	BM122020	BM122020	RPCT-24-3
C 387	18	2.2	516	12	BJ187611	BJ187611		C 460	18	2.2	567	9	AL826015	AL826015	AL826015
C 388	18	2.2	516	13	BU436829	BU436829	604146037	C 461	18	2.2	567	12	BM306915	BM306915	ECESTE45
C 389	18	2.2	517	14	CA561929	CA561929	K0294A11-	C 462	18	2.2	568	10	BF056935	BF056935	7k12b03.x
C 390	18	2.2	519	28	AQ307477	AQ307477	HS.2178.B	C 463	18	2.2	569	13	BQ459150	BQ459150	
C 391	18	2.2	522	9	AU199487	AU199487	AU39487	C 464	18	2.2	570	2	HSM084665	HSM084665	
C 392	18	2.2	522	12	BM779227	BM779227	EST589802	C 465	18	2.2	570	14	CB828105	CB828105	LjNEST829
C 393	18	2.2	523	10	BF010308	BF010308	689B08-.Y	C 466	18	2.2	571	12	BM110196	BM110196	
C 394	18	2.2	523	28	AZ521543	AZ521543	RPCT-11-1	C 467	18	2.2	572	9	AM092048	AM092048	
C 395	18	2.2	524	28	AQ405957	AQ405957	HS.5043.B	C 468	18	2.2	572	10	BM586822	BM586822	EST488591
C 396	18	2.2	525	13	BQ092028	BQ092028	F436a11.x	C 469	18	2.2	572	12	BI209643	BI209643	EST57683
C 397	18	2.2	525	13	BQ486645	BQ486645	1091046F0	C 470	18	2.2	573	12	BM109643	BM109643	
C 398	18	2.2	526	10	BE801303	BE801303	BT12D04.Y	C 471	18	2.2	576	9	AL673516	AL673516	AL673516
C 399	18	2.2	527	9	AJ483793	AJ483793	EST249664	C 472	18	2.2	576	10	BF269664	BF269664	CA_Bb000
C 400	18	2.2	527	9	AJ435139	AJ435139		C 473	18	2.2	576	29	B2922966	B2922966	EST240_115
C 401	18	2.2	527	9	AJ924374	AJ924374		C 474	18	2.2	577	13	BQ117706	BQ117706	
C 402	18	2.2	528	9	AJ442181	AJ442181		C 475	18	2.2	577	14	CB911681	CB911681	VVD134E07
C 403	18	2.2	528	12	BJ136279	BJ136279	8a58c07.Y	C 476	18	2.2	579	28	BS2876	BS2876	CIT-HSP-200
C 404	18	2.2	528	13	BQ394710	BQ394710	NISC.ng10	C 477	18	2.2	585	9	AV930167	AV930167	AV930167
C 405	18	2.2	529	9	AV984794	AV984794		C 478	18	2.2	586	9	AV673631	AV673631	AV673631
C 406	18	2.2	529	28	AQ050519	AQ050519	RPCT-11-4	C 479	18	2.2	586	13	BM207891	BM207891	AV673919
C 407	18	2.2	530	28	AZ051525	AZ051525		C 480	18	2.2	586	14	CB214972	CB214972	NISC.np01
C 408	18	2.2	532	12	BM821998	BM821998	K-EST0091	C 481	18	2.2	586	14	CB214972	CB214972	BF0505467
C 409	18	2.2	534	9	AA907680	AA907680	om11d05.8	C 482	18	2.2	587	10	BF505467	BF505467	
C 410	18	2.2	534	13	BM312212	BM312212		C 483	18	2.2	587	14	H76020	H76020	
C 411	18	2.2	537	9	AM034438	AM034438	EST278009	C 484	18	2.2	587	28	BM373955	BM373955	
C 412	18	2.2	537	12	BI501785	BI501785	rm05a07.Y	C 485	18	2.2	588	29	DR36C14T	DR36C14T	
C 413	18	2.2	537	28	BH251469	BH251469	SAIK.0116	C 486	18	2.2	589	14	C92459	C92459	
C 414	18	2.2	538	9	AM380436	AM380436	RC1-HT026	C 487	18	2.2	589	28	BH732284	BH732284	BOMKR31TR
C 415	18	2.2	538	9	AM652669	AM652669	100884.MA	C 488	18	2.2	590	14	CA642897	CA642897	wreln.pk0
C 416	18	2.2	538	12	BJ151272	BJ151272		C 489	18	2.2	593	12	BM308450	BM308450	eak4c05.
C 417	18	2.2	539	13	BU780829	BU780829	STJEGF05	C 490	18	2.2	594	14	CD524887	CD524887	
C 418	18	2.2	539	12	BJ173111	BJ173111		C 491	18	2.2	595	14	CA061295	CA061295	gsl1c08.Y
C 419	18	2.2	540	10	BG095629	BG095629	EST460184	C 492	18	2.2	596	13	BU991053	BU991053	HD05N047
C 420	18	2.2	541	9	AV954918	AV954918		C 493	18	2.2	597	29	B2919954	B2919954	CH240.69M
C 421	18	2.2	542	12	BJ422298	BJ422298		C 494	18	2.2	598	9	AU239598	AU239598	AU239598
C 422	18	2.2	545	10	BE449474	BE449474	EST356233	C 495	18	2.2	598	9	AV782821	AV782821	AV782821
C 423	18	2.2	547	10	BF223155	BF223155	7q29905.x	C 496	18	2.2	598	12	BJ207305	BJ207305	BM207305
C 424	18	2.2	547	28	AQ377108	AQ377108	RPCT11-15	C 497	18	2.2	600	29	AG243020	AG243020	Lotub.jap
C 425	18	2.2	548	9	AJ928240	AJ928240	w095910.x	C 498	18	2.2	601	29	BZ287137	BZ287137	SAIK.0205
C 426	18	2.2	548	9	AL388633	AL388633		C 499	18	2.2	602	12	BI921729	BI921729	EST541632
C 427	18	2.2	548	9	AU199767	AU199767		C 500	18	2.2	605	9	AV983760	AV983760	AV983760
C 428	18	2.2	548	12	BI501551	BI501551	rm01h04.Y	C 501	18	2.2	605	10	BF490540	BF490540	AT27037.5
C 429	18	2.2	548	12	BJ401734	BJ401734		C 502	18	2.2	606	12	BI206106	BI206106	EST524146
C 430	18	2.2	548	14	CA920260	CA920260	EST637978	C 503	18	2.2	607	9	AJ1727005	AJ1727005	
C 431	18	2.2	549	10	BE802821	BE802821	gr45a11.Y	C 504	18	2.2	607	12	BJ139130	BJ139130	
C 432	18	2.2	549	28	AQ754369	AQ754369	HS.5394.B	C 505	18	2.2	607	28	AQ575557	AQ575557	nbx00087F
C 433	18	2.2	550	9	AL681945	AL681945	tx50h08.x	C 506	18	2.2	608	28	BM396693	BM396693	AG-ND-156
C 434	18	2.2	550	14	CD525084	CD525084	kw12h05.Y	C 507	18	2.2	609	28	AQ3774805	AQ3774805	
C 435	18	2.2	551	14	CD533246	CD533246	ECESTee66	C 508	18	2.2	611	9	AV916297	AV916297	AV916297
C 436	18	2.2	553	9	AU213122	AU213122		C 509	18	2.2	611	13	BM231727	BM231727	
C 437	18	2.2	553	12	BJ462537	BJ462537		C 510	18	2.2	613	10	BF474342	BF474342	WHE0842.H
C 438	18	2.2	555	9	AV993047	AV993047		C 511	18	2.2	613	28	B29896	B29896	TI8018TFC.T
C 439	18	2.2	555	12	BJ137784	BJ137784		C 512	18	2.2	616	13	BQ465773	BQ465773	
C 440	18	2.2	556	12	BI441085	BI441085	1c80f08.Y	C 513	18	2.2	617	9	AJ436555	AJ436555	
C 441	18	2.2	556	12	BJ104239	BJ104239		C 514	18	2.2	617	9	AM241826	AM241826	xn75902.x
C 442	18	2.2	557	13	BQ621724	BQ621724	TVEST017.	C 515	18	2.2	617	13	BM120526	BM120526	BM120526

516	18	2.2	619	9	AL789754	AL789754	AL789754	589	18	2.2	670	13	BU704382	BU704382	UI-M-F00-
C 517	18	2.2	619	12	BU127685	BU127685	BU127685	590	18	2.2	670	29	BM241632	BM241632	AG160117
C 518	18	2.2	620	9	AUI98645	AUI98645	AUI98645	591	18	2.2	671	29	AG160117	AG160117	AG178230
C 519	18	2.2	620	9	AV653165	AV653165	AV653165	592	18	2.2	671	29	AG178230	AG178230	BE463316
C 520	18	2.2	620	9	AV955269	AV955269	AV955269	593	18	2.2	672	10	BE463316	BE463316	EST354539
C 521	18	2.2	620	10	BF274780	BF274780	BF274780	594	18	2.2	673	29	AG180815	AG180815	AG180815
C 522	18	2.2	620	10	BF274780	BF274780	BF274780	595	18	2.2	674	10	BF493683	BF493683	AG180815
C 523	18	2.2	620	12	BU120231	BU120231	BU120231	596	18	2.2	674	28	BH964683	BH964683	AG180815
C 524	18	2.2	621	9	BU121633	BU121633	BU121633	597	18	2.2	675	12	BU142788	BU142788	AG180815
C 525	18	2.2	621	9	AU236546	AU236546	AU236546	598	18	2.2	676	9	AU061517	AU061517	AG180815
C 526	18	2.2	624	12	BU399097	BU399097	BU399097	599	18	2.2	676	12	BU133660	BU133660	AG180815
C 527	18	2.2	624	13	BQ459883	BQ459883	BQ459883	600	18	2.2	677	28	AZ205562	AZ205562	AG180815
C 528	18	2.2	626	9	AL791618	AL791618	AL791618	601	18	2.2	677	28	AJ499122	AJ499122	AG180815
C 529	18	2.2	627	14	CA769437	CA769437	CA769437	602	18	2.2	678	13	BM197828	BM197828	AG180815
C 530	18	2.2	627	18	AZ799526	AZ799526	AZ799526	603	18	2.2	681	14	CA051839	CA051839	AG180815
C 531	18	2.2	627	29	BZ469119	BZ469119	BZ469119	604	18	2.2	682	9	AV995675	AV995675	AG180815
C 532	18	2.2	628	10	BE256176	BE256176	BE256176	605	18	2.2	682	28	BZ004733	BZ004733	AG180815
C 533	18	2.2	628	10	BE256176	BE256176	BE256176	606	18	2.2	682	28	BZ078073	BZ078073	AG180815
C 534	18	2.2	628	14	CA393428	CA393428	CA393428	607	18	2.2	683	10	BG107797	BG107797	AG180815
C 535	18	2.2	628	14	CB831422	CB831422	CB831422	608	18	2.2	683	13	BM248039	BM248039	AG180815
C 536	18	2.2	629	9	AV672242	AV672242	AV672242	609	18	2.2	687	13	BM045921	BM045921	AG180815
C 537	18	2.2	629	12	BU027843	BU027843	BU027843	610	18	2.2	688	29	BZ911623	BZ911623	AG180815
C 538	18	2.2	630	29	AG078749	AG078749	AG078749	611	18	2.2	688	29	AG138058	AG138058	AG180815
C 539	18	2.2	631	14	AV826881	AV826881	AV826881	612	18	2.2	689	9	AI807903	AI807903	AG180815
C 540	18	2.2	631	14	CA661684	CA661684	CA661684	613	18	2.2	689	10	BE474742	BE474742	AG180815
C 541	18	2.2	633	9	AV999195	AV999195	AV999195	614	18	2.2	691	28	AZ504907	AZ504907	AG180815
C 542	18	2.2	633	13	BM218457	BM218457	BM218457	615	18	2.2	692	13	BM774149	BM774149	AG180815
C 543	18	2.2	633	14	CD002176	CD002176	CD002176	616	18	2.2	692	13	BM262824	BM262824	AG180815
C 544	18	2.2	634	29	AG164093	AG164093	AG164093	617	18	2.2	695	12	BU135258	BU135258	AG180815
C 545	18	2.2	635	9	AM607414	AM607414	AM607414	618	18	2.2	695	13	BM248142	BM248142	AG180815
C 546	18	2.2	636	9	AV821970	AV821970	AV821970	619	18	2.2	695	29	BZ977744	BZ977744	AG180815
C 547	18	2.2	636	9	AV995722	AV995722	AV995722	620	18	2.2	696	9	AM442794	AM442794	AG180815
C 548	18	2.2	636	13	BM260938	BM260938	BM260938	621	18	2.2	697	13	BU001959	BU001959	AG180815
C 549	18	2.2	636	14	CA754564	CA754564	CA754564	622	18	2.2	697	13	BM269302	BM269302	AG180815
C 550	18	2.2	637	13	BQ459766	BQ459766	BQ459766	623	18	2.2	698	9	AV782563	AV782563	AG180815
C 551	18	2.2	637	28	AQ782652	AQ782652	AQ782652	624	18	2.2	699	13	BM271965	BM271965	AG180815
C 552	18	2.2	638	13	BM032680	BM032680	BM032680	625	18	2.2	699	28	AZ046173	AZ046173	AG180815
C 553	18	2.2	638	28	BH918943	BH918943	BH918943	626	18	2.2	700	29	AG113915	AG113915	AG180815
C 554	18	2.2	639	13	BM191728	BM191728	BM191728	627	18	2.2	701	9	AJ424817	AJ424817	AG180815
C 555	18	2.2	639	13	BM216726	BM216726	BM216726	628	18	2.2	701	13	BM262348	BM262348	AG180815
C 556	18	2.2	641	12	BM288188	BM288188	BM288188	629	18	2.2	701	29	CNS034XL	CNS034XL	AG180815
C 557	18	2.2	642	12	BU466551	BU466551	BU466551	630	18	2.2	702	13	BM256377	BM256377	AG180815
C 558	18	2.2	642	13	BU412242	BU412242	BU412242	631	18	2.2	706	10	BM398204	BM398204	AG180815
C 559	18	2.2	642	14	BY723214	BY723214	BY723214	632	18	2.2	706	13	BM237499	BM237499	AG180815
C 560	18	2.2	642	14	W20061	W20061	W20061	633	18	2.2	706	28	BH982591	BH982591	AG180815
C 561	18	2.2	644	10	BF504110	BF504110	BF504110	634	18	2.2	707	28	BH939941	BH939941	AG180815
C 562	18	2.2	644	13	BM304350	BM304350	BM304350	635	18	2.2	708	12	BI091417	BI091417	AG180815
C 563	18	2.2	645	13	BM238619	BM238619	BM238619	636	18	2.2	709	13	BM110696	BM110696	AG180815
C 564	18	2.2	646	9	AV675177	AV675177	AV675177	637	18	2.2	710	28	BZ089277	BZ089277	AG180815
C 565	18	2.2	647	9	AV960785	AV960785	AV960785	638	18	2.2	710	28	BE883002	BE883002	AG180815
C 566	18	2.2	648	12	BP018121	BP018121	BP018121	639	18	2.2	711	10	BE997770	BE997770	AG180815
C 567	18	2.2	648	13	BM257571	BM257571	BM257571	640	18	2.2	711	10	BE997770	BE997770	AG180815
C 568	18	2.2	649	13	BM295198	BM295198	BM295198	641	18	2.2	711	14	CD002814	CD002814	AG180815
C 569	18	2.2	649	29	BX206921	BX206921	BX206921	642	18	2.2	712	9	AV722542	AV722542	AG180815
C 570	18	2.2	651	29	AG176317	AG176317	AG176317	643	18	2.2	713	3	AV676031	AV676031	AG180815
C 571	18	2.2	652	14	CA008433	CA008433	CA008433	644	18	2.2	713	9	BM248714	BM248714	AG180815
C 572	18	2.2	652	28	AZ570406	AZ570406	AZ570406	645	18	2.2	713	13	BM271334	BM271334	AG180815
C 573	18	2.2	654	13	BM053784	BM053784	BM053784	646	18	2.2	714	13	BM294756	BM294756	AG180815
C 574	18	2.2	654	13	BM237189	BM237189	BM237189	647	18	2.2	715	14	CB459480	CB459480	AG180815
C 575	18	2.2	656	9	AM686107	AM686107	AM686107	648	18	2.2	716	12	BU148563	BU148563	AG180815
C 576	18	2.2	656	13	BM158295	BM158295	BM158295	649	18	2.2	716	13	BU148563	BU148563	AG180815
C 577	18	2.2	656	18	BM284142	BM284142	BM284142	650	18	2.2	716	13	BM259422	BM259422	AG180815
C 578	18	2.2	659	29	CC231632	CC231632	CC231632	651	18	2.2	716	28	BM259422	BM259422	AG180815
C 579	18	2.2	660	13	BQ407243	BQ407243	BQ407243	652	18	2.2	716	28	BH37786	BH37786	AG180815
C 580	18	2.2	660	29	CC066362	CC066362	CC066362	653	18	2.2	716	28	BH37786	BH37786	AG180815
C 581	18	2.2	660	29	CC066362	CC066362	CC066362	654	18	2.2	716	28	BH37786	BH37786	AG180815
C 582	18	2.2	661	29	AG128082	AG128082	AG128082	655	18	2.2	717	28	BH022794	BH022794	AG180815
C 583	18	2.2	662	9	AV823002	AV823002	AV823002	656	18	2.2	718	13	BM273977	BM273977	AG180815
C 584	18	2.2	662	13	BM268011	BM268011	BM268011	657	18	2.2	721	13	BM282142	BM282142	AG180815
C 585	18	2.2	662	29	AG176247	AG176247	AG176247	658	18	2.2	722	12	BU12484	BU12484	AG180815
C 586	18	2.2	663	13	BU091680	BU091680	BU091680	659	18	2.2	722	12	BU430792	BU430792	AG180815
C 587	18	2.2	667	13	BM238000	BM238000	BM238000	660	18	2.2	722	29	BZ821116	BZ821116	AG180815
C 588	18	2.2	667	13	BM270881	BM270881	BM270881	661	18	2.2	723	28	AZ423050	AZ423050	AG180815

662	18	2.2	726	13	BW287311	BW287311	BW287311	735	18	2.2	848	28	AZ683865	AZ683865	ENTMA77TR
663	18	2.2	726	14	CA429633	CA429633	CA429633	736	18	2.2	850	13	BO893155	BO893155	BO893155
664	18	2.2	727	29	B2392990	B2392990	B2392990	737	18	2.2	852	28	BH503206	BH503206	BOGCV76TR
665	18	2.2	729	13	BW293081	BW293081	BW293081	738	18	2.2	854	28	AZ693608	AZ693608	ENTLJ76TR
666	18	2.2	730	12	B1955451	B1955451	B1955451	739	18	2.2	857	29	CNS0070R	CNS0070R	AL066257
667	18	2.2	730	28	BH129948	BH129948	BH129948	740	18	2.2	858	13	BH338224	BH338224	BOJ531978
668	18	2.2	732	13	BW200633	BW200633	BW200633	741	18	2.2	863	29	CC404872	CC404872	CC404872
669	18	2.2	733	13	BW253829	BW253829	BW253829	742	18	2.2	867	13	BU364233	BU364233	BU364233
670	18	2.2	734	9	AJ392783	AJ392783	AJ392783	743	18	2.2	868	29	CNS02EM5	CNS02EM5	AL194220
671	18	2.2	737	28	BH573799	BH573799	BH573799	744	18	2.2	869	28	BH159853	BH159853	AL194220
672	18	2.2	738	14	CD576302	CD576302	CD576302	745	18	2.2	872	29	CC407768	CC407768	CC407768
673	18	2.2	741	13	BW271034	BW271034	BW271034	746	18	2.2	874	14	CBS66552	CBS66552	CC407768
674	18	2.2	741	10	BG648392	BG648392	BG648392	747	18	2.2	881	28	AZ684482	AZ684482	CC407768
675	18	2.2	744	28	AO937906	AO937906	AO937906	748	18	2.2	884	13	BU770589	BU770589	BU770589
676	18	2.2	745	28	BSM078981	BSM078981	BSM078981	749	18	2.2	891	29	CC243426	CC243426	CC243426
677	18	2.2	746	13	BW158266	BW158266	BW158266	750	18	2.2	896	28	AZ547221	AZ547221	CC243426
678	18	2.2	748	12	BW999256	BW999256	BW999256	751	18	2.2	897	29	AGJ31302	AGJ31302	CC243426
679	18	2.2	749	9	AV723004	AV723004	AV723004	752	18	2.2	897	29	CNS05F0R	CNS05F0R	CC243426
680	18	2.2	758	9	AM053909	AM053909	AM053909	753	18	2.2	898	14	CBS61920	CBS61920	CC243426
681	18	2.2	758	28	AO858640	AO858640	AO858640	754	18	2.2	902	14	BO900676	BO900676	CC243426
682	18	2.2	760	13	BW216077	BW216077	BW216077	755	18	2.2	903	13	BH38724	BH38724	CC243426
683	18	2.2	761	12	BG829742	BG829742	BG829742	756	18	2.2	906	28	AZ570270	AZ570270	CC243426
684	18	2.2	762	13	BW257439	BW257439	BW257439	757	18	2.2	911	28	AGJ31302	AGJ31302	CC243426
685	18	2.2	764	13	BW208020	BW208020	BW208020	758	18	2.2	911	29	CNS013P8	CNS013P8	CC243426
686	18	2.2	765	29	BX131386	BX131386	BX131386	759	18	2.2	916	29	CNS000J0	CNS000J0	CC243426
687	18	2.2	766	12	BJ576165	BJ576165	BJ576165	760	18	2.2	917	29	CC243426	CC243426	CC243426
688	18	2.2	766	13	BW278969	BW278969	BW278969	761	18	2.2	918	28	AZ548818	AZ548818	CC243426
689	18	2.2	768	12	BJ377640	BJ377640	BJ377640	762	18	2.2	923	29	CNS07A0H	CNS07A0H	CC243426
690	18	2.2	768	28	BH990450	BH990450	BH990450	763	18	2.2	934	14	CD514661	CD514661	CC243426
691	18	2.2	769	13	BW211382	BW211382	BW211382	764	18	2.2	936	29	BZ821111	BZ821111	CC243426
692	18	2.2	770	28	BH378709	BH378709	BH378709	765	18	2.2	936	29	CC243426	CC243426	CC243426
693	18	2.2	775	10	BG213955	BG213955	BG213955	766	18	2.2	940	10	CG025555	CG025555	CC243426
694	18	2.2	775	10	BF270022	BF270022	BF270022	767	18	2.2	942	9	AL666573	AL666573	CC243426
695	18	2.2	776	13	BW122598	BW122598	BW122598	768	18	2.2	942	29	CNS04B0G	CNS04B0G	CC243426
696	18	2.2	777	10	BE033858	BE033858	BE033858	769	18	2.2	946	13	BU149189	BU149189	CC243426
697	18	2.2	777	10	BE052743	BE052743	BE052743	770	18	2.2	948	29	CNS03GHC	CNS03GHC	CC243426
698	18	2.2	778	29	BZ489859	BZ489859	BZ489859	771	18	2.2	954	28	AZ692619	AZ692619	CC243426
699	18	2.2	780	29	BZ789381	BZ789381	BZ789381	772	18	2.2	963	29	CNS01HTM	CNS01HTM	CC243426
700	18	2.2	781	13	BW569489	BW569489	BW569489	773	18	2.2	964	29	CNS00U8S	CNS00U8S	CC243426
701	18	2.2	783	9	AJ514167	AJ514167	AJ514167	774	18	2.2	965	10	BE053656	BE053656	CC243426
702	18	2.2	783	28	BH418711	BH418711	BH418711	775	18	2.2	973	29	BZ481562	BZ481562	CC243426
703	18	2.2	784	28	BH387640	BH387640	BH387640	776	18	2.2	973	29	CNS07C02	CNS07C02	CC243426
704	18	2.2	788	14	CD493271	CD493271	CD493271	777	18	2.2	975	29	CC198406	CC198406	CC243426
705	18	2.2	788	28	BH703175	BH703175	BH703175	778	18	2.2	981	29	CNS005SY	CNS005SY	CC243426
706	18	2.2	788	28	BZ808612	BZ808612	BZ808612	779	18	2.2	984	29	CNS00F8H	CNS00F8H	CC243426
707	18	2.2	790	13	BW237843	BW237843	BW237843	780	18	2.2	984	29	CNS06DBH	CNS06DBH	CC243426
708	18	2.2	790	29	BX178880	BX178880	BX178880	781	18	2.2	993	10	BE053389	BE053389	CC243426
709	18	2.2	791	14	CB662594	CB662594	CB662594	782	18	2.2	996	13	BX325165	BX325165	CC243426
710	18	2.2	792	14	CD493689	CD493689	CD493689	783	18	2.2	999	29	CNS00268	CNS00268	CC243426
711	18	2.2	793	14	CD493689	CD493689	CD493689	784	18	2.2	1006	13	BX372820	BX372820	CC243426
712	18	2.2	797	13	BO137572	BO137572	BO137572	785	18	2.2	1016	28	BH022511	BH022511	CC243426
713	18	2.2	798	12	BI253742	BI253742	BI253742	786	18	2.2	1029	13	BX436064	BX436064	CC243426
714	18	2.2	799	28	AZ892473	AZ892473	AZ892473	787	18	2.2	1032	28	BH245711	BH245711	CC243426
715	18	2.2	800	9	AL667603	AL667603	AL667603	788	18	2.2	1038	29	CNS00388	CNS00388	CC243426
716	18	2.2	801	28	BH099353	BH099353	BH099353	789	18	2.2	1039	13	BU186284	BU186284	CC243426
717	18	2.2	803	12	BI648541	BI648541	BI648541	790	18	2.2	1081	29	CNS054YG	CNS054YG	CC243426
718	18	2.2	803	29	BZ523896	BZ523896	BZ523896	791	18	2.2	1091	12	BM806740	BM806740	CC243426
719	18	2.2	810	28	BH655089	BH655089	BH655089	792	18	2.2	1091	12	CNS03VYH	CNS03VYH	CC243426
720	18	2.2	811	29	CC425335	CC425335	CC425335	793	18	2.2	1096	29	CC270901	CC270901	CC243426
721	18	2.2	813	13	BU224551	BU224551	BU224551	794	18	2.2	1097	29	CNS0035N	CNS0035N	CC243426
722	18	2.2	813	29	CC425335	CC425335	CC425335	795	18	2.2	1101	29	CNS00GNH	CNS00GNH	CC243426
723	18	2.2	816	28	AO253275	AO253275	AO253275	796	18	2.2	1101	29	CNS00GNH	CNS00GNH	CC243426
724	18	2.2	817	29	BZ466591	BZ466591	BZ466591	797	18	2.2	1106	10	BP138480	BP138480	CC243426
725	18	2.2	817	29	BX209382	BX209382	BX209382	798	18	2.2	1165	9	AL533835	AL533835	CC243426
726	18	2.2	830	12	BI648541	BI648541	BI648541	799	18	2.2	1166	14	BM473684	BM473684	CC243426
727	18	2.2	830	28	BH263025	BH263025	BH263025	800	18	2.2	1172	29	CC262245	CC262245	CC243426
728	18	2.2	832	10	BM447198	BM447198	BM447198	801	18	2.2	1176	29	CC262245	CC262245	CC243426
729	18	2.2	835	13	BG447198	BG447198	BG447198	802	18	2.2	1284	14	CC498563	CC498563	CC243426
730	18	2.2	843	14	BY762094	BY762094	BY762094	803	18	2.2	1390	29	CC191816	CC191816	CC243426
731	18	2.2	844	10	BF623152	BF623152	BF623152	804	18	2.2	1390	29	CC191816	CC191816	CC243426
732	18	2.2	845	28	BG865296	BG865296	BG865296	805	18	2.2	1559	12	BI524414	BI524414	CC243426
733	18	2.2	845	28	BZ113385	BZ113385	BZ113385	806	18	2.2	1662	10	BE963522	BE963522	CC243426
734	18	2.2	847	28	AZ691322	AZ691322	AZ691322	807	18	2.2	1718	11	AK049146	AK049146	CC243426

C 808	18	2.2	4743	11	BC040213	C 881	17	2.1	275	9	A1553208	A1553208 v117e04.y
C 809	17	2.1	63	29	BX218898	C 882	17	2.1	275	10	BE069634	BE069634 RC2-BT038
C 810	17	2.1	98	29	BZ767070	C 883	17	2.1	275	13	B0030228	B0030228 CHU1D04.
C 811	17	2.1	107	9	AA120480	C 884	17	2.1	278	4	BX519439	BX519439 RZPD Mbs
C 812	17	2.1	144	9	AM278282	C 885	17	2.1	279	9	A1068275	A1068275 EST20966
C 813	17	2.1	151	9	A1286990	C 886	17	2.1	279	9	AV025436	AV025436 AV025436
C 814	17	2.1	161	12	BJ521706	C 887	17	2.1	280	9	A1197646	A1197646 ue46C03.x
C 815	17	2.1	163	9	AV357659	C 888	17	2.1	280	14	CD343432	CD343432 ECEStee70
C 816	17	2.1	170	14	CB891677	C 889	17	2.1	282	14	CB079817	CB079817 hp77c06.b
C 817	17	2.1	173	10	BG136975	C 890	17	2.1	283	9	AA783544	AA783544 c6a07a1.f
C 818	17	2.1	178	10	BE028449	C 891	17	2.1	283	9	AV831322	AV831322 AV831322
C 819	17	2.1	180	9	AM163919	C 892	17	2.1	286	9	AA785933	AA785933 17c06a1.f
C 820	17	2.1	180	12	BJ521697	C 893	17	2.1	286	28	AQ267579	AQ267579 RPTC11-69
C 821	17	2.1	182	9	A1210101	C 894	17	2.1	287	9	AM587330	AM587330 EST318953
C 822	17	2.1	182	29	BZ355072	C 895	17	2.1	288	14	U46354	U46354 HSU46354 Hu
C 823	17	2.1	184	10	BE002469	C 896	17	2.1	289	28	AQ092155	AQ092155 HS 3013.A
C 824	17	2.1	185	9	AL958347	C 897	17	2.1	290	10	BF842266	BF842266 MRI-HT113
C 825	17	2.1	185	10	BG211825	C 898	17	2.1	291	9	AM690683	AM690683 NF032E105
C 826	17	2.1	186	29	AG219571	C 899	17	2.1	291	9	AA518560	AA518560 v117e04.x
C 827	17	2.1	187	13	BQ975530	C 900	17	2.1	292	9	AM128543	AM128543 fe17c01.y
C 828	17	2.1	189	28	AZ010778	C 901	17	2.1	292	10	BG221285	BG221285 RST41089
C 829	17	2.1	192	9	AM215611	C 902	17	2.1	292	10	BR282179	BR282179 BR282179
C 830	17	2.1	195	10	BE028694	C 903	17	2.1	293	10	BR449467	BR449467 BR449467
C 831	17	2.1	199	10	BE069638	C 904	17	2.1	294	10	BR262490	BR262490 BR262490
C 832	17	2.1	200	9	A1302283	C 905	17	2.1	294	10	BB333174	BB333174 BB333174
C 833	17	2.1	200	9	AV150644	C 906	17	2.1	295	9	AV114467	AV114467 AV114467
C 834	17	2.1	206	9	AV152513	C 907	17	2.1	295	9	AA331977	AA331977 EST35863
C 835	17	2.1	207	9	AU264273	C 908	17	2.1	295	10	BB269065	BB269065 BB269065
C 836	17	2.1	216	9	AM784047	C 909	17	2.1	297	9	AA772412	AA772412 a144e12.b
C 837	17	2.1	217	10	AM993236	C 910	17	2.1	300	14	C54172	C54172 C54172 Yu11
C 838	17	2.1	219	14	CB277349	C 911	17	2.1	300	28	BB750316	BB750316 SMLK 0377
C 839	17	2.1	220	10	BG364330	C 912	17	2.1	301	9	AV213216	AV213216 AV213216
C 840	17	2.1	220	12	BM321889	C 913	17	2.1	302	9	AA966247	AA966247 v7a07a1.f
C 841	17	2.1	222	28	BH192797	C 914	17	2.1	302	9	A1209513	A1209513 b0e01a1.f
C 842	17	2.1	222	28	BH192797	C 915	17	2.1	302	9	AV152299	AV152299 AV152299
C 843	17	2.1	223	10	BG139003	C 916	17	2.1	303	9	AA685125	AA685125 EST106314
C 844	17	2.1	226	9	A1022470	C 917	17	2.1	304	14	D67546	D67546 CELK053HXF
C 845	17	2.1	227	9	BB017189	C 918	17	2.1	307	14	T94084	T94084 ye28h09.a1
C 846	17	2.1	232	9	A1481958	C 919	17	2.1	309	28	AZ888554	AZ888554 RPTC-24-1
C 847	17	2.1	232	14	T01084	C 920	17	2.1	310	10	BG196341	BG196341 RST15556
C 848	17	2.1	235	10	BG045362	C 921	17	2.1	310	10	BB6826	BB6826 RPTC11-26B2
C 849	17	2.1	235	12	BM493266	C 922	17	2.1	315	12	B1072539	B1072539 C077P880
C 850	17	2.1	236	14	T37505	C 923	17	2.1	315	29	BZ764315	BZ764315 SMLK 1244
C 851	17	2.1	238	9	AV076703	C 924	17	2.1	315	29	BB764316	BB764316 SMLK 1244
C 852	17	2.1	240	9	AU073295	C 925	17	2.1	317	9	A1328137	A1328137 x8e06f1.f
C 853	17	2.1	242	10	BB488718	C 926	17	2.1	317	9	AW023756	AW023756 dF5H03.y
C 854	17	2.1	243	10	AL958555	C 927	17	2.1	318	9	A1211408	A1211408 poa09a1.f
C 855	17	2.1	243	14	CB820512	C 928	17	2.1	318	12	BM429776	BM429776 1Du021C1.
C 856	17	2.1	244	10	BG459225	C 929	17	2.1	318	13	B0574403	B0574403 PA Ba000
C 857	17	2.1	246	10	BE028434	C 930	17	2.1	318	28	AZ808105	AZ808105 2M0071021
C 858	17	2.1	247	10	BE028719	C 931	17	2.1	320	10	BG198880	BG198880 RST18150
C 859	17	2.1	250	12	BM306113	C 932	17	2.1	322	13	BQ334117	BQ334117 MK0-WT012
C 860	17	2.1	250	13	BU096591	C 933	17	2.1	323	9	AA068647	AA068647 mm63b10.x
C 861	17	2.1	252	9	AA966498	C 934	17	2.1	324	9	A1486044	A1486044 EST244365
C 862	17	2.1	253	14	D43467	C 935	17	2.1	324	12	BM062041	BM062041 KS01037H0
C 863	17	2.1	253	14	D43467	C 936	17	2.1	325	14	CD346120	CD346120 ECEStee95
C 864	17	2.1	254	14	R64964	C 937	17	2.1	325	14	T43918	T43918 7181 Lambda
C 865	17	2.1	255	10	AA126772	C 938	17	2.1	329	10	BE376894	BE376894 RCO-TN008
C 866	17	2.1	255	10	BE463554	C 939	17	2.1	330	10	BE069556	BE069556 RC2-BT038
C 867	17	2.1	255	13	BX472226	C 940	17	2.1	331	13	BQ3188504	BQ3188504 AGENCOURT
C 868	17	2.1	255	28	AZ306479	C 941	17	2.1	331	14	CA918956	CA918956 EST616674
C 869	17	2.1	259	14	CD346633	C 942	17	2.1	331	14	CD347382	CD347382 ECESteeF08
C 870	17	2.1	263	9	AA442499	C 943	17	2.1	334	9	AA969079	AA969079 op55c04.B
C 871	17	2.1	263	14	F37139	C 944	17	2.1	334	14	CD343358	CD343358 ECEStee67
C 872	17	2.1	263	28	AZ641084	C 945	17	2.1	334	14	CD346107	CD346107 ECEStee95
C 873	17	2.1	268	12	B1972704	C 946	17	2.1	335	29	BX206000	BX206000 Danilo rer
C 874	17	2.1	268	12	B1972704	C 947	17	2.1	338	9	AU110304	AU110304 AU110304
C 875	17	2.1	269	9	AV223650	C 948	17	2.1	338	12	B1021824	B1021824 RC5-WT025
C 876	17	2.1	271	13	BU029678	C 949	17	2.1	339	13	BQ998142	BQ998142 OCG18H13.
C 877	17	2.1	272	10	BB503855	C 950	17	2.1	339	13	BY034704	BY034704 BY034704
C 878	17	2.1	273	12	B1000142	C 951	17	2.1	339	28	AQ086696	AQ086696 nbe0b0029H
C 879	17	2.1	274	12	BM321725	C 952	17	2.1	340	10	BE605382	BE605382 f125g10.y
C 880	17	2.1	274	29	CNS00VBS	C 953	17	2.1	342	9	AA466346	AA466346 l02ag92 N

```

954 17 2.1 342 28 BH848928
955 17 2.1 343 13 BQ30342
956 17 2.1 343 14 CA773634
957 17 2.1 348 9 A1327801
958 17 2.1 349 9 A1756835
959 17 2.1 349 14 CD343122
960 17 2.1 349 14 CD343329
961 17 2.1 349 14 CD343508
962 17 2.1 350 10 BF068967
963 17 2.1 351 9 AV980471
964 17 2.1 352 9 AV734422
965 17 2.1 352 9 AG251887
966 17 2.1 353 28 AQ284792
967 17 2.1 354 13 BY019647
968 17 2.1 355 9 AM278294
969 17 2.1 357 9 AV749215
970 17 2.1 358 9 AA354835
971 17 2.1 358 29 BX200597
972 17 2.1 360 9 AA067964
973 17 2.1 360 14 C60170
974 17 2.1 360 14 C62948
975 17 2.1 360 14 D68498
976 17 2.1 360 28 AQ706592
977 17 2.1 361 10 BF818546
978 17 2.1 365 9 AA892410
979 17 2.1 365 14 CB212688
980 17 2.1 366 9 AA096774
981 17 2.1 366 9 A1523355
982 17 2.1 366 14 CD051842
983 17 2.1 366 14 CD346548
984 17 2.1 367 13 BX259340
985 17 2.1 368 9 AM822358
986 17 2.1 369 9 A1445971
987 17 2.1 370 13 BY397242
988 17 2.1 371 9 AL046024
989 17 2.1 372 10 BF802868
990 17 2.1 375 9 AA755639
991 17 2.1 375 9 AM207920
992 17 2.1 376 9 AA922170
993 17 2.1 376 12 BM257506
994 17 2.1 377 9 AV561261
995 17 2.1 377 9 AM520548
996 17 2.1 377 13 BU040325
997 17 2.1 377 14 CA999847
998 17 2.1 378 12 B1979133
999 17 2.1 379 13 BQ791801
1000 17 2.1 379 28 AQ300048

```

## ALIGNMENTS

```

RESULT 1          371 bp  mRNA  linear  EST 20-DEC-2000
BF646896          clone NF07D01EC1F012 Ellicted cell culture Medicago truncatula cDNA
LOCUS             NF07D01EC1F012 5' mRNA sequence.
DEFINITION
ACCESSION          BF646896
VERSION            BF646896.1 GI:11912026
KEYWORDS
SOURCE
ORGANISM           Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; euroids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
1 (bases 1 to 371)
Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Imman,J.T., Weller,J.W. and May,G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation -
Center for Medicago Genomics Research
Unpublished
JOURNAL
COMMENT            Contact: Dixon RA

```

```

FEATURES
  source
    1..371
      /organism="Medicago truncatula"
      /mol_type="mRNA"
      /db_xref="taxon:3880"
      /clone="NF07D01EC"
      /tissue_type="Cell cultures derived from root tissues"
      /dev_stage="Cell suspensions were subcultured every 14
days. Cells were induced six days after subculture"
      /note="Vector: Lambda Zap; Cells were induced with yeast
cell wall extracts equivalent to 50ug/ml glucose in the
final concentration. Samples were taken at 0.5, 1, 12 and
24 hours after induction. Equal amounts of RNA from each
time point were pooled and used for mRNA isolation."
    126 a 56 c 83 g 106 t

```

```

BASE COUNT
ORIGIN
Query Match 4.2%; Score 35; DB 10; Length 371;
Best Local Similarity 100.0%; Pred.No. 2.7e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Cy 151 AGTTGATGATGAGTGCTGATTAATCTAAGCC 185
Db 248 AGTTGATGATGAGTGCTGATTAATCTAAGCC 282

```

```

RESULT 2          671 bp  mRNA  linear  EST 16-MAR-2001
BG447824          clone NF103E12EC1F097 Ellicted cell culture Medicago truncatula cDNA
LOCUS             NF103E12EC1F097 5' mRNA sequence.
DEFINITION
ACCESSION          BG447824
VERSION            BG447824.1 GI:13366605
KEYWORDS
SOURCE
ORGANISM           Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; euroids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
1 (bases 1 to 671)
Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Imman,J.T., Weller,J.W. and May,G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation -
Center for Medicago Genomics Research
Unpublished
JOURNAL
COMMENT            Contact: Dixon RA

```

```

FEATURES
  source
    1..671
      /organism="Medicago truncatula"
      /mol_type="mRNA"
      /db_xref="taxon:3880"
      /clone="NF103E12EC"
      /tissue_type="Cell cultures derived from root tissues"

```

/dev stage="Cell suspensions were subcultured every 14 days. Cells were induced six days after subculture"  
/clone lib="Elicited cell culture"  
/note="Vector: Lambda Zap. Cells were induced with yeast cell wall extracts equivalent to 50ug/ml glucose in the final concentration. Samples were taken at 0.5, 1, 12 and 24 hours after induction. Equal amounts of RNA from each time point were pooled and used for RNA isolation."

BASE COUNT 235 a 122 c 128 g 183 t 3 others

ORIGIN

Query Match 4.2%; Score 35; DB 10; Length 671;  
Best Local Similarity 100.0%; Pred. No. 3.4e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 AGTTGAGATGAGTGGCTGATATCTAAGGCC 185  
Db 248 AGTTGAGATGAGTGGCTGATATCTAAGGCC 282

RESULT 3  
BE462229 335 bp mRNA linear EST 18-MAY-2001  
LOCUS EST32453 tomato flower buds 0-3 mm, Cornell University  
DEFINITION Lycopersicon esculentum cDNA clone cTOA12C3, mRNA sequence.  
ACCESSION BE462229  
VERSION BE462229.1 GI:9507998  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 335)  
van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksey,S.D.  
Generation of ESTs from tomato flower tissue, 0-3 mm buds  
Unpublished  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES  
source  
1..335  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4081"  
/db\_xref="TA496"  
/clone="cTOA12C3"  
/tissue\_type="flower"  
/dev stage="0-3mm buds"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Tanksey; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

BASE COUNT 109 a 54 c 77 g 95 t

ORIGIN

Query Match 4.1%; Score 34; DB 10; Length 335;  
Best Local Similarity 100.0%; Pred. No. 9.1e-07;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 175 TATCTAAGGCCACATATCAAGAGGCTGACTTG 208  
Db 185 TATCTAAGGCCACATATCAAGAGGCTGACTTG 218

RESULT 4

AM626121 418 bp mRNA linear EST 18-MAY-2001  
LOCUS EST320028 tomato radicle, 5 d post-imbibition, Cornell University  
DEFINITION Lycopersicon esculentum cDNA clone cLE218K18 5', mRNA sequence.  
ACCESSION AM626121  
VERSION AM626121.1 GI:7339148  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 418)  
van der Hoeven,R.S., Garvin,D., Matern,A.L., Holt,I.E., Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksey,S.D.  
Generation of ESTs from tomato radicle tissue (etiolated)  
Unpublished  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES  
source  
1..418  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4081"  
/db\_xref="TA496"  
/clone="cLE218K18"  
/tissue\_type="radicle"  
/dev stage="seedlings 5 days post-imbibition"  
/clone lib="tomato radicle, 5 d post-imbibition, Cornell University"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Tanksey; Tissue supplied by Dave Garvin (USDA-ARS, Ithaca, NY 14850)."  
1 (bases 1 to 418)  
van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksey,S.D.  
Generation of ESTs from tomato flower tissue, 0-3 mm buds  
Unpublished  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA

BASE COUNT 128 a 79 c 91 g 119 t 1 others

ORIGIN

Query Match 4.1%; Score 34; DB 9; Length 418;  
Best Local Similarity 100.0%; Pred. No. 9.9e-07;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 175 TATCTAAGGCCACATATCAAGAGGCTGACTTG 208  
Db 177 TATCTAAGGCCACATATCAAGAGGCTGACTTG 210

RESULT 5  
BE462282 490 bp mRNA linear EST 18-MAY-2001  
LOCUS EST324546 tomato flower buds 0-3 mm, Cornell University  
DEFINITION Lycopersicon esculentum cDNA clone cTOA12C2, mRNA sequence.  
ACCESSION BE462282  
VERSION BE462282.1 GI:9508051  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 490)  
van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksey,S.D.  
Generation of ESTs from tomato flower tissue, 0-3 mm buds  
Unpublished  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA



Email: <http://www.genome.clemson.edu/orders/index.html>

## FEATURES

5 prime sequence.  
Location/Qualifiers  
1. 490

/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA96"  
/db\_xref="taxon:4081"  
/clone="CTOA12C2"  
/tissue\_type="flower"  
/dev\_stage="0-3mm buds"  
/clone\_id="tomato flower buds 0-3 mm, Cornell University"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Tankley; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA96). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

BASE COUNT 156 a 93 c 105 g 136 t

## ORIGIN

Query Match 4.1%; Score 34; DB 10; Length 490;  
Best Local Similarity 100.0%; Pred. No. 1e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 TATCTAAGGCCACATATCAAGAGTGCCTTTG 208  
|||||  
177 TATCTAAGGCCACATATCAAGAGTGCCTTTG 210

## RESULT 6

BF519256

LOCUS 245 bp mRNA linear EST 08-DEC-2000  
DEFINITION B519256.1 DSIL Medicago truncatula cDNA clone pDSIL-20W9, mRNA

ACCESSION B519256  
VERSION B519256.1 GI:11607939

KEYWORDS EST.  
SOURCE Medicago truncatula (barrel medic)

## ORGANISM

Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;  
Medicago.

1 (bases 1 to 245)

REFERENCE  
AUTHORS Fedorova, M., Pierson, B. L., Samac, D. A., Vance, C. P., Gantt, G. S., Peng,  
H., Ellis, L., Town, C. D., Bowman, C. L., Craven, M. B., Hansen, T. S.,  
Holt, I. E. and Fraser, C. M.  
ESTs from leaves of Medicago truncatula after inoculation with  
Colletotrichum trifolii

## TITLE

## JOURNAL

## COMMENT

Unpublished  
Contact: Deborah A. Samac  
Department of Plant Pathology  
University of Minnesota  
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA  
Tel: 612 625 1243  
Fax: 651 649 5058  
Email: debby@pnucini.crl.umn.edu  
TIGR sequence name: M7PCE77TK  
More information is available at: <http://chryste.tamu.edu/medicago>  
Seq primer: SKmod (CTA GAA CTA GCG GAT CC).

## FEATURES

SOURCE

1. 245  
Location/Qualifiers  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/cultivar="genotype A17"  
/db\_xref="taxon:3880"  
/clone="pDSIL-20W9"  
/tissue\_type="leaves infected with Colletotrichum  
trifolii"  
/dev\_stage="cotyledons and primary leaves harvested 5 and  
8 days after inoculation with Colletotrichum trifolii"  
/lab\_host="E. coli strain XL0R"

/clone\_id="DSIL"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; cDNA was prepared from polyA+ enriched RNA from  
cotyledons and primary leaves harvested 5 and 8 days after  
inoculation with Colletotrichum trifolii. The cDNA was  
directionally ligated into the Uni-ZAP XR vector from  
Stratagene and packaged using Gigapack III Gold packaging  
extracts. Plasmids containing cDNA inserts were excised  
from the recombinant lambda-ZAP phage using Ex-Assist  
helper phage and propagated in XL0R cells. Note: EST may  
be of fungal origin."

BASE COUNT 74 a 32 c 66 g 73 t

## ORIGIN

Query Match 3.4%; Score 28; DB 10; Length 245;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 GATTGAGGTGGCTGAATTATCTAAGGCC 185  
|||||  
172 GATTGAGGTGGCTGAATTATCTAAGGCC 199

## RESULT 7

BF635572

LOCUS 367 bp mRNA linear EST 19-DEC-2000  
DEFINITION NF104H01DTNF1014 Drought Medicago truncatula cDNA clone NF104H01DT  
5', mRNA sequence.

ACCESSION BF635572  
VERSION BF635572.1 GI:11899730

KEYWORDS EST.  
SOURCE Medicago truncatula (barrel medic)

## ORGANISM

Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;  
Medicago.

1 (bases 1 to 367)

REFERENCE  
AUTHORS Torres-Jerez, I., Scott, A. D., Harris, A. R., Gonzales, R. A., Bell, C. J.,  
Flores, H. R., Iman, J. T., Weller, J. W. and May, G. D.  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula drought library

## TITLE

## JOURNAL

## COMMENT

Unpublished  
Contact: May GD  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org  
Insert length: 367 Std Error: 0.00  
Plate: 104 row: H column: 01  
Seq primer: TCACACGGAACACGCTATGAC.

## FEATURES

SOURCE

1. 367  
Location/Qualifiers  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/db\_xref="taxon:3880"  
/clone="NF104H01DT"  
/tissue\_type="Plantlets"  
/dev\_stage="Pooled timepoints"  
/clone\_id="Drought"  
/note="Vector: Lambda Zap; Contains a mixture of entire  
plantlets harvested in a series of days-post-watering  
timepoints."

BASE COUNT 142 a 50 c 85 g 86 t 4 others

## ORIGIN

Query Match 3.3%; Score 27; DB 10; Length 367;  
Best Local Similarity 100.0%; Pred. No. 0.0058;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 TGTCGAAAAGTTGATGAGGTGG 168



```

Db      107  TGTGGAAAAGTTGATGATGAGTGG 133
|||||
RESULT 8
AM980341
LOCUS
DEFINITION AM980341 526 bp mRNA linear EST 07-SEP-2000
sequence.
ACCESSION EST391494 GVN Medicago truncatula cDNA clone pGVN-30E12, mRNA
AM980341
VERSION
KEYWORDS
SOURCE
ORGANISM Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 526)
Pedrova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,J.S., Peng
,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
Holt,I.E. and Fraser,C.M.
ESTs from one month old nitrogen-fixing root nodules of Medicago
truncatula
JOURNAL
COMMENT
Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651-649-5058
Email: vance004@maroon.tc.umn.edu
University of Minnesota name:M256175e
TIGR sequence name:MTCBH30TK
More information is available at: http://chrysis.tamu.edu/medicago
Seq primer: Skmod (CTA gaa cta gtc gat cc).
Location/Qualifiers
1. 526
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pGVN-30E12"
/tissue_type="N2-fixing root nodules"
/dev_stage="effective root nodules harvested one month
post inoculation with Sinorhizobium meliloti"
/lab_host="E. coli strain XLOLR"
/clone_lib="GVN"
/notes="Vector: Bluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
effective root nodules harvested one month post
inoculation with Sinorhizobium meliloti. The cDNA was
directionally ligated into the Uni-ZAP XR vector from
stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-ZAP phage using Ex-Assist
helper phage and propagated in XLOLR cells."
BASE COUNT 183 a 75 c 108 g 160 t
ORIGIN
Query Match 3.3%; Score 27; DB 10; Length 526;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 151 AGTTGATTGAGTGGCTGAATTAT 177
|||||
Db 383 AGTTGATTGAGTGGCTGAATTAT 409
|||||
RESULT 9
BF006562 526 bp mRNA linear EST 06-OCT-2000
LOCUS
DEFINITION BF006562 DSLC Medicago truncatula cDNA clone pDSL-42B13, mRNA
EST435060 DSLC Medicago truncatula cDNA clone pDSL-42B13, mRNA

```

```

sequence.
ACCESSION BF006562
VERSION BF006562.1 GI:10706837
KEYWORDS
SOURCE
ORGANISM Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 526)
Pedrova,M., Pierson,B.L., Samac,D.A., Gantt,J.S., Vance,C.P.,
Gonzalez,M.B. and Ellis,L.
ESTs from Medicago truncatula leaves and cotyledons
JOURNAL
COMMENT
Contact: Deborah A. Samac
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debys@puccini.crl.umn.edu
University of Minnesota name: M275600e TIGR sequence name:
MTLB507TK More information is available at:
http://chrysis.tamu.edu/medicago
Seq primer: Skmod (CTA gaa cta gtc gat cc).
Location/Qualifiers
1. 526
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pDSL-42B13"
/tissue_type="leaves and cotyledons"
/dev_stage="mixture of cotyledons from five days old
plants and leaves obtained from two weeks old plants"
/lab_host="E. coli strain SOLR"
/clone_lib="DSL"
/notes="Vector: Bluescript SK +/-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA from the
mixture of cotyledons of five days old plants and leaves
of two weeks old plants. The cDNA was directionally
ligated into the Uni-ZAP XR vector from Stratagene and
packaged using Gigapack III Gold packaging extracts.
Plasmids containing cDNA inserts were excised from the
recombinant lambda-ZAP phage using Ex-Assist helper phage
and propagated in SOLR cells."
BASE COUNT 191 a 98 c 110 g 127 t
ORIGIN
Query Match 3.3%; Score 27; DB 10; Length 526;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 151 AGTTGATTGAGTGGCTGAATTAT 177
|||||
Db 238 AGTTGATTGAGTGGCTGAATTAT 264
|||||
RESULT 10
BE124179 549 bp mRNA linear EST 07-SEP-2000
LOCUS
DEFINITION BE124179 DSLC Medicago truncatula cDNA clone pDSL-13B12, mRNA
sequence.
ACCESSION BE124179
VERSION BE124179.1 GI:8528736
KEYWORDS
SOURCE
ORGANISM Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

```

REFERENCE 1 (bases 1 to 549)  
 AUTHORS Fedorova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng  
 TITLE H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,  
 Holt,I.B., and Fraser,C.M.  
 ESTs from leaves of *Medicago truncatula* after inoculation with  
*Colletotrichum trifolii*  
 JOURNAL Unpublished  
 COMMENT Contact: Deborah A. Samac  
 Department of Plant Pathology  
 University of Minnesota  
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA  
 Tel: 612 625 1243  
 Fax: 651 649 5058  
 Email: debby@puccini.crl.umn.edu  
 Minnesota sequence name: M262137e  
 TIGR sequence name: MTFBH307K8  
 More information is available at:  
<http://chryste.tamu.edu/medicago>  
 Seq primer: Skmod (CTA gaa CTA gtc gat CC).  
 Location/Qualifiers  
 1..549  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /cultivar="genotype A17"  
 /db\_xref="taxon:3880"  
 /clone="pDSIL-13E12"  
 /tissue="pDSIL-13E12"  
 /issue\_type="leaves infected with *Colletotrichum trifolii*"  
 /dev\_stage="cotyledons and primary leaves harvested 5 and  
 8 days after inoculation with *Colletotrichum trifolii*"  
 /lab\_host="E. coli strain XLOLR"  
 /clone\_lib="DSIL"  
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
 XhoI; cDNA was prepared from polyA+ enriched RNA from  
 cotyledons and primary leaves harvested 5 and 8 days after  
 inoculation with *Colletotrichum trifolii*. The cDNA was  
 directionally ligated into the Uni-ZAP XR vector from  
 Stratagene and packaged using Gigapack III Gold packaging  
 extracts. Plasmids containing cDNA inserts were excised  
 from the recombinant lambda-ZAP phage using Ex-Assist  
 helper phage and propagated in XLOLR cells. Note: EST may  
 be of fungal origin."
 

BASE COUNT 194 a 80 c 113 g 162 t

ORIGIN

Query Match 3.3%; Score 27; DB 10; Length 549;  
 Best Local Similarity 100.0%; Pred. No. 0.0068;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 AGTTGATTGAGTGCTGAATTAT 177  
 ||||||||||||||||||||||||  
 Db 381 AGTTGATTGAGTGCTGAATTAT 407

RESULT 11  
 LOCUS BG588067 585 bp mRNA linear EST 12-APR-2001  
 DEFINITION EST1483876 MHRP- *Medicago truncatula* cDNA clone pMHRP-41C11, mRNA  
 sequence.  
 ACCESSION BG588067  
 VERSION BG588067.1 GI:13606207  
 KEYWORDS EST.  
 ORGANISM *Medicago truncatula* (barrel medic)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.  
 1 (bases 1 to 585)  
 Harrison,M.U., Liu,J., Town,C.D., Van Aken,S., Uteback,T., Cho,J.  
 and Fraser,C.M.  
 TITLE ESTs from phosphate-starved roots of *Medicago truncatula*, 2001  
 JOURNAL Unpublished

COMMENT Contact: Harrison M.U.  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73401  
 Tel: 580-223-5810  
 Fax: 580-221-7380  
 Email: mjharrison@noble.org  
 The Samuel Roberts Noble Foundation: N266740e TIGR sequence name:  
 MTHA187KB More information is available at:  
<http://www.medicago.org>  
 Seq primer: Skmod (CTA gaa CTA gtc gat CC).  
 Location/Qualifiers  
 1..585  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /cultivar="A17"  
 /db\_xref="taxon:3880"  
 /clone="pMHRP-41C11"  
 /tissue\_type="roots"  
 /dev\_stage="phosphate-starved"  
 /lab\_host="XLOLR"  
 /clone\_lib="MHRP-"  
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
 XhoI; At the trifoliolate stage, M. truncatula plants were  
 transplanted to phosphate-free sand and grown for a  
 further 30 days. During this period, they were fertilized  
 twice weekly with 1/2 Hoaglands solutions containing 20uM  
 potassium phosphate. cDNA was prepared from polyA+  
 enriched RNA. The cDNA was directionally ligated into the  
 UniZap XR vector from Stratagene and packaged using  
 Gigapack III Gold packaging extracts. Plasmids containing  
 cDNA inserts were excised from the recombinant lambda-ZAP  
 phage using Ex-assist helper phage and propagated in  
 XLOLR cells."
 

BASE COUNT 194 a 106 c 130 g 155 t

ORIGIN

Query Match 3.3%; Score 27; DB 10; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 0.0069;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 AGTTGATTGAGTGCTGAATTAT 177  
 ||||||||||||||||||||||||  
 Db 167 AGTTGATTGAGTGCTGAATTAT 193

RESULT 12  
 LOCUS BG647939 618 bp mRNA linear EST 24-APR-2001  
 DEFINITION EST509558 HOGA *Medicago truncatula* cDNA clone pHOGA-18G24 5' end,  
 mRNA sequence.  
 ACCESSION BG647939  
 VERSION BG647939.1 GI:13783051  
 KEYWORDS EST.  
 SOURCE *Medicago truncatula*  
 ORGANISM *Medicago truncatula* (barrel medic)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.  
 1 (bases 1 to 618)  
 Hahn,M.G., Ojane-Reuns,T., Samac,D., Town,C.D., Van Aken,S.,  
 Uteback,T., Cho,J., and Fraser,C.M.  
 TITLE ESTs from roots of *Medicago truncatula* treated with  
 oligogalacturonides of DP 6-20  
 JOURNAL Unpublished  
 COMMENT Contact: Michael G. Hahn  
 Complex Carbohydrate Research Center  
 University of Georgia  
 220 Riverbend Road, Athens, GA 30602-4712, USA  
 Tel: 706-542-4457  
 Fax: 706-542-4412  
 Email: hahn@ccrc.uga.edu

FEATURES  
source

G391545e TIGR sequence name: MTMCF48TK More information is available at: [www.medicago.org](http://www.medicago.org)  
Seq primer: Skmod (CTA gaa cta gtc gat cc).  
Location/Qualifiers

1.618  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/cultivar="A17"  
/db\_xref="taxon:3880"  
/clone="pHOGA-16G24"  
/tissue\_type="3 day old seedling roots"  
/dev\_stage="24 hours after treatment in the dark at 26 C with 0.5 mg/ml oligogalacturonides (DP 6-20) in the presence of 100 ug/ml Gentamicin"  
/lab\_host="XLOLR"  
/clone\_lib="HOGA"  
/note="Vector: Bluescript SK-; Site 1: EcoRI, Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-aseist helper phage and propagated in SOLR cells."

BASE COUNT 211 a 118 c 117 g 172 t

Query Match 3.3%; Score 27; DB 10; Length 618;  
Best Local Similarity 100.0%; Pred. No. 0.0071;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 ATTGAGTGGCTGATATCTAAGGCC 185  
|||||  
Db 87 ATTGAGTGGCTGATATCTAAGGCC 113

RESULT 13  
BG647226 651 bp mRNA linear EST 24-APR-2001  
LOCUS EST508845 HOGA Medicago truncatula cDNA clone pHOGA-16E11 5' end,  
DEFINITION mRNA sequence.  
ACCESSION BG647226 GI:13782338  
VERSION BG647226.1  
KEYWORDS EST.  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula

REFERENCE  
AUTHORS Hahn,M.G., Ojansen-Reuhs,T., Samac,D., Town,C.D., Van Aken,S.,  
TITLE Uteback,T., Cho,J. and Frazer,C.M.  
JOURNAL ESTs from roots of Medicago truncatula created with  
COMMENT oligogalacturonides of DP 6-20  
Unpublished  
Contact: Michael G. Hahn  
Complex Carbohydrate Research Center  
University of Georgia  
220 Riverbend Road, Athens, GA 30602-4712, USA  
Tel: 706-542-4457  
Fax: 706-542-4412  
Email: hahn@ccrc.uga.edu  
G390832e TIGR sequence name: MTMBW30TK More information is  
available at: [www.medicago.org](http://www.medicago.org)  
Seq primer: Skmod (CTA gaa cta gtc gat cc).  
Location/Qualifiers

FEATURES  
source

1..651  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/cultivar="A17"  
/db\_xref="taxon:3880"  
/clone="pHOGA-16E11"  
/tissue\_type="3 day old seedling roots"

/dev\_stage="24 hours after treatment in the dark at 26 C with 0.5 mg/ml oligogalacturonides (DP 6-20) in the presence of 100 ug/ml Gentamicin"  
/lab\_host="XLOLR"  
/clone\_lib="HOGA"  
/note="Vector: Bluescript SK-; Site 1: EcoRI, Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-aseist helper phage and propagated in SOLR cells."

BASE COUNT 223 a 124 c 132 g 172 t

Query Match 3.3%; Score 27; DB 10; Length 651;  
Best Local Similarity 100.0%; Pred. No. 0.0072;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 AGTTGAGATTGAGTGGCTGAATTAT 177  
|||||  
Db 232 AGTTGAGATTGAGTGGCTGAATTAT 258

RESULT 14  
BU830456 781 bp mRNA linear EST 15-OCT-2002  
LOCUS BU830456  
DEFINITION T008508 Populus apical shoot cDNA library Populus tremula x Populus tremuloides cDNA 5 prime, mRNA sequence.  
ACCESSION BU830456 GI:24007683  
VERSION BU830456.1  
KEYWORDS EST.  
SOURCE Populus tremula x Populus tremuloides  
ORGANISM Populus tremula x Populus tremuloides

REFERENCE  
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
TITLE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
JOURNAL ; eucosids I; Malvaginales; Salicaceae; Populus.  
COMMENT 1 (bases 1 to 781)  
The poplar tree trancriptome: Analysis of expressed sequence tags  
from multiple libraries  
Unpublished  
Contact: BHALERAO RUPALI R.  
Umea Plant Science Center  
Department of Plant Physiology  
University of Umea, 901 87 Umea, Sweden  
Tel: +46 90 786 5279  
Fax: +46 90 786 6676  
Email: rupali.bhalerao@plantphys.umu.se.  
Location/Qualifiers

FEATURES  
source

1..781  
/organism="Populus tremula x Populus tremuloides"  
/mol\_type="mRNA"  
/db\_xref="taxon:47664"  
/tissue\_type="apical shoot"  
/clone\_lib="Populus apical shoot cDNA library"

BASE COUNT 270 a 147 c 183 g 181 t

Query Match 3.2%; Score 26; DB 11; Length 781;  
Best Local Similarity 100.0%; Pred. No. 0.027;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 CATTAAGCTCTTACGCAACAGATGTC 266  
|||||  
Db 279 CATTAAGCTCTTACGCAACAGATGTC 304

RESULT 15  
BU876538 308 bp mRNA linear EST 16-OCT-2002  
LOCUS BU876538/c  
DEFINITION V021H11 Populus flower cDNA library Populus balsamifera subsp.  
trichocarpa cDNA 5 prime, mRNA sequence.

ACCESSION BU876538  
 VERSION BU876538.1 GI:24068062  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM *Populus balsamifera* subsp. *trichocarpa*  
*Populus balsamifera* subsp. *trichocarpa*  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids I; Malpighiales; Salicaceae; *Populus*.  
 1 (bases 1 to 308)  
 UNEBERG, P., BHALERAO, R. R., JANSON, S. and STERKY, F.  
 The poplar tree transcriptome: Analysis of expressed sequence tags  
 from multiple libraries  
 Unpublished  
 JOURNAL  
 COMMENT Contact: BHALERAO RUPALI R.  
 Umea Plant Science Center  
 Department of Plant Physiology  
 University of Umea, 901 87 Umea, Sweden  
 Tel: +46 90 786 5279  
 Fax: +46 90 786 6676  
 Email: rupali.bhalerao@plantphys.umu.se.  
 FEATURES  
 source  
 1..308  
 location/Qualifiers  
 /organism="Populus balsamifera subsp. trichocarpa"  
 /mol\_type="mRNA"  
 /sub\_species="trichocarpa"  
 /db\_xref="taxon:3694"  
 /clone\_lib="Populus flower cDNA library"  
 /note="Organ: flower"  
 BASE COUNT 88 a 75 c 59 g 86 t  
 ORIGIN  
 Query Match 3.0%; Score 25; DB 13; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 0.066;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 590 ATGACGATATTGAAGAAGATGAGA 614  
 ||||||||||||||||||  
 DB 139 ATGACGATATTGAAGAAGATGAGA 115

Search completed: January 30, 2004, 10:47:17  
 Job time : 2028 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 29, 2004, 20:04:01 ; Search time 41 Seconds  
(without alignments)  
1060.758 Million cell updates/sec

Title: US-10-033-190-2

Perfect score: 1477  
Sequence: 1 MNSTMSNLGVRKSWTDEE.....QQGVNDPFSAEIDLWNLLD 274

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*  
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1477	100.0	274	23	ABB81626
2	542.5	36.7	22	AAE01932	Tomato anthocyanin
3	542.5	36.7	22	AAE01932	Arabidopsis thalia
4	538	36.4	23	AAU93177	Arabidopsis transc
5	538	36.2	24	AAU93174	A. thaliana Produc
6	532.5	36.1	22	AAE01916	Amino acid sequenc
7	532.5	36.1	23	AAU92969	Arabidopsis thalia
8	532.5	36.1	24	AAE00046	Arabidopsis thalia
9	527	35.7	23	AAU97535	A. thaliana Produc

10	524	35.5	212	23	AAU93155	Arabidopsis transc
11	518.5	35.1	209	22	AAE01933	Arabidopsis thalia
12	518.5	35.1	209	23	AAU93176	Arabidopsis transc
13	510	34.5	211	22	AAE68355	Amino acid sequenc
14	422	28.6	302	22	AAE82483	Cotton transcript
15	420	28.4	391	21	AAE33304	Pinus radiata tran
16	415	28.1	205	21	AAE33293	Eucalyptus grandis
17	408	27.6	258	21	AAE29949	Arabidopsis thalia
18	408	27.6	259	21	AAE29948	Arabidopsis thalia
19	408	27.6	259	22	AAE02486	Arabidopsis thalia
20	408	27.6	278	21	AAE29947	Arabidopsis thalia
21	403	27.3	246	21	AAE30442	Arabidopsis thalia
22	403	27.3	257	21	AAE30441	Cotton transcript
23	402.5	27.3	226	22	AAE82482	Arabidopsis transc
24	401	27.1	203	23	AAU93131	Pinus radiata tran
25	400	27.1	148	21	AAE33044	Pinus radiata tran
26	399.5	27.0	176	21	AAE33283	Arabidopsis transc
27	397	26.9	342	23	AAU93088	Myb-related transc
28	396.5	26.8	302	23	ABJ10412	Myb-related transc
29	393	26.6	120	23	ABJ10409	Eucalyptus grandis
30	388.5	26.3	153	21	AAE33239	Myb-related transc
31	387.5	26.2	253	23	ABJ10420	Pinus radiata tran
32	387.5	26.2	255	21	AAE33306	Pinus radiata tran
33	385	26.1	246	23	ABJ10422	Myb-related transc
34	379.5	25.7	258	21	AAE08529	Arabidopsis thalia
35	377.5	25.6	236	21	AAE02485	Arabidopsis thalia
36	377.5	25.6	236	21	AAE02485	Arabidopsis thalia
37	377.5	25.6	236	21	AAE02485	Arabidopsis thalia
38	377.5	25.6	236	21	AAE02485	Arabidopsis thalia
39	377.5	25.6	236	21	AAE02485	Arabidopsis thalia
40	377.5	25.6	236	21	AAE02485	Arabidopsis thalia
41	375	25.4	280	21	AAE02485	Arabidopsis thalia
42	375	25.4	280	22	AAE68353	Amino acid sequenc
43	374	25.3	371	22	AAE10416	Myb-related transc
44	374	25.3	371	22	AAE01913	Arabidopsis thalia
45	374	25.3	371	23	AAU92995	Arabidopsis transc

#### ALIGNMENTS

RESULT 1  
ABB81626  
ID ABB81626 standard; Protein; 274 AA.  
AC ABB81626;  
XX 24-SEP-2002 (first entry)  
DT 24-SEP-2002 (first entry)  
XX Tomato anthocyanin 1 (ANT1) protein SEQ ID NO:2.  
XX DE Tomato anthocyanin 1 (ANT1) protein SEQ ID NO:2.  
XX KM Tomato; ANT1; anthocyanin 1; Lycopersicon esculentum; plant; colour.  
XX OS Lycopersicon esculentum.  
XX PN W0200255658-A2.  
XX PD 18-JUL-2002.  
XX PF 29-OCT-2001; 2001MO-US50638.  
XX PR 30-OCT-2000; 2000US-244685P.  
XX PA (EXEL-) EXELIXIS PLANT SCI INC.  
XX PI Connors K, Mathews HV, Liu A;  
XX DR WPI: 2002-557819/59.  
XX DR N-PSDB; ABQ73046.  
XX PT New isolated polynucleotide derived from tomato, useful for producing an Anthocyanin 1 phenotype in plants, particularly for modifying e.g. leaf color, flower color or fruit color in plants



28-FEB-2002.

22-AUG-2001; 2001WO-US26189.

23-AUG-2000; 2000US-227439P.

16-NOV-2000; 2000US-0713994.

16-APR-2001; 2001US-0837944.

(MEND-) MENDEL BIOTECHNOLOGY INC.

(PIG/) PILGRIM M.

(CREE/) CREELMAN R.

(DUBE/) DUBELL A J.

(HEAR/) HEARD J.

(JIANG/) JIANG C.

(KEDD/) KEDDIE J.

(ADAM/) ADAM L.

(RATC/) RATCLIFF O.

(REUB/) REUBER J L.

(RIEC/) RIECHMANN J L.

(YUGG/) YU G.

(PINE/) PINEDA O.

Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J, Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;

WPI; 2002-292022/33.

N-PSDB; ABK65363.

An isolated or recombinant polynucleotide used to produce a transgenic plant -

Claim 40; Page 880-881; 941pp; English.

The invention relates to 1 of 232 isolated or recombinant polynucleotides encoding an Arabidopsis thaliana transcription factor, their variants, complements, fragments, or related polynucleotide with 31% to 95% sequence identity, where the plant possesses an altered trait as compared to a wild-type or reference plant, or the plant exhibits an altered phenotype as compared to a wild-type or reference plant, or the plant exhibits ecotypic expression or altered expression of one or more genes associated with a plant trait as compared to a wild plant. Also included are a transgenic plant comprising the polynucleotides, a computer readable medium having stored sequence information, and identifying a homologue sequence from a database comprising a plurality of known plant sequences comprising inputting sequence information selected from one of 464 fully defined sequences given in the specification. The isolated or recombinant polynucleotide is used for producing a plant having a modified trait, the method comprising selecting a polynucleotide that encodes a polypeptide or an antisense nucleic acid, inserting the polynucleotide or antisense nucleic acid into an expression vector, introducing the vector into a plant or a cell of a plant to overexpress the polypeptide or antisense nucleic acid, thereby producing a modified plant, and selecting for a modified trait (e.g. increased production of agriculturally useful proteins or metabolic chemicals, pest tolerance, environmentally stress response (e.g. drought), microbial disease resistance, herbicide resistance, seed and fruit yield, growth rate, leaf and flower senescence and many other traits listed in the specification). The present sequence is one of the 232 proteins which are A. thaliana transcription factors.

Sequence 246 AA;

Query Match 36.7%; Score 542.5; DB 23; Length 246;

Best Local Similarity 52.4%; Pred. No. 1e-43;

Matches 109; Conservative 26; Mismatches 52; Indels 21; Gaps 4;

7 SSLGVRKGSWTDEEDFLRKIDIKYGEKWHLPVRAGLNRKRCSCGLRWMLNYLRPIKR 66

4 SPKGRKGTCTTDEEDILKQCIDIKYGEKWHLPVRAGLNRKRCSCGLRWMLNYLRPIKR 63

67 GDFEODEVDLILRLKILGNRWSLIAGRLPGRANDVKNYWNLTLLKLN---TTKIVP 122

64 GKLCSDVDLIVLRLKILGNRWSLIAGRLPGRANDVKNYWNLTLLKLNBERCKTKMI- 122

123 REKINKCGEISTKBIKIPORRRKFSITKMYNT-----NNVILDEEH 167

123 NKNITSHPTSSAQKIDVLRKPRPSFSDKNSCNVNLKPKVDVVPRLHGLNNNYVCSSIT 182

168 C-KEIISEKQTPDASMDNDPFWMINILE 194

183 CNKDEQDKLNLINLLDGDNMWSEILLE 210

RESULT 4

AAU75734

AAU75734 standard; Protein; 248 AA.

AAU75734;

08-MAY-2002 (first entry)

A. thaliana production of anthocyanin pigment 1 (PAP1) protein.

PAP1; production of anthocyanin pigment; MYB-like transcription factor; transgenic; plant; phenylpropanoid; stress; light stress; water stress; pH stress; temperature stress; heavy metal stress; pathogen attack; infection; wounding; nutrient deficiency; herbivory; plant colouration; stress sensitisation; chromosome 1.

Arabidopsis thaliana.

WO200200902-A2.

03-JAN-2002.

21-JUN-2001; 2001WO-US19734.

23-JUN-2000; 2000US-0603244.

05-JUL-2000; 2000US-0610185.

(SALK ) SALK INST BIOLOGICAL STUDIES.

(ROBE-) ROBERTS NOBLE FOUND INC SAMMUEL.

Borevitz J, Xia Y, Lamb CJ, Dixon RA;

WPI; 2002-16443/21.

N-PSDB; ABK14221.

New isolated PAP1 or PAP2 gene, useful for increasing pigmentation in plants, as reporter genes for analysing expression pattern of promoter of interest, and to increase flux through phenylpropanoid pathway -

Claim 9; Page 26-27; 29pp; English.

This invention relates to the nucleotide and protein sequences of novel Arabidopsis thaliana production of anthocyanin pigment (PAP) genes PAP1 or PAP2. The genes for both PAP1 and PAP2 have been mapped to chromosome 1. PAP1 and PAP2 are MYB-like transcription factors that regulate the production of anthocyanin pigment. The nucleotide sequence of the invention is useful for screening for PAP gene expression, it is also useful for detecting promoter activity within a plant cell. The nucleotide sequence is also useful for screening for plant cell (e.g., A. thaliana) transformation. The transformed plant cells are then used to produce transgenic plants with tissue-specific altered colour expression. The nucleotide sequence is also useful for enhancing accumulation of phenylpropanoid product in a plant. A transgenic plant is useful for detecting a stress condition such as light stress, water stress, pH stress, temperature stress, heavy metal stress, pathogen attack or infection, wounding, nutrient deficiency, herbivory, or abnormal hormone levels in a plant. PAP1 or PAP2 overexpression can also be useful in developing or enhancing desirable colouration in ornamental plants for commercial sales, and for sensitising plants to stress. Introduction of PAP1 or PAP2 sensitises plants to respond more quickly to a stress condition in ways that are physiologically significant, as well as in ways of ornamental significance. The present sequence represents the A. thaliana PAP1

CC protein sequence of the invention.

XX Sequence 248 AA;

Query Match 36.4%; Score 538; DB 23; Length 248;  
Best Local Similarity 38.9%; Pred. No. 2.8e-43;  
Matches 115; Conservative 47; Mismatches 48; Indels 86; Gaps 10;

QY 7 SSGVKGKSWTDEEDFLRKCIDKYGKXHLVPIRAGLNCRKSCRLRLNTLRLPHIKR 66  
DB 4 SSKGLRGKAWTDEEDSLRQICINKYEGKXHVPLRAGLNCRKSCRLRLNTLRLPHIKR 63  
QY 67 GDFODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNL-----LRKL 115  
DB 64 GKLSNDEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLSSKHSCKSKMKK 123  
QY 116 NTKKI-----VPREK--INNKGEIST--KIEIIPQRRKYESTKNTNN 159  
DB 124 DITPIPTPALKNVYKPRPSFTVNNDCNHLNAPRYDVNP-----CLGLINNNVCDNS 179  
QY 160 VIIDEEHCKEIISEKQTPDASNDVDPWMTNLLENCDIIEDEEVYINYEKTLTSLH 219  
DB 180 IITNKDKKKQDLV-----NLIDGNMMLKRF-----LEESQEV----- 213  
QY 220 EEISPPINIGSGMSMOGQISHEWGEFSLNLPFMQGVONDDPSAID-LMNLID 274  
DB 214 -----DLVPEATTEKQ-----DTLAFVDQLMSLFD 241

RESULT 5  
AAB68356  
ID AAB68356 standard; Protein; 249 AA.

XX AAB68356;

DT 09-JUL-2001 (first entry)

DE Amino acid sequence of MYB transcription factor AtMYB90.

KM Stress tolerance; myloblastosis transcription factor; heat tolerance;  
KW MYB transcription factor; AtMYB60; AtMYB74; AtMYB75; transgenic plant;  
KM phenylpropanoid biosynthesis.

XX Arabidopsis sp.

XX MO200132002-A1.

XX 10-MAY-2001.

XX 06-NOV-2000; 2000WO-US30503.

XX 05-NOV-1999; 99US-0163579.

XX 23-OCT-2000; 2000US-0693855.

XX (BADI ) BASF CORP.

XX Tonelli C;

XX WPI, 2001-316365/33.

XX N-PSDB; AAF85191.

PT Nucleic acid molecules encoding plant stress tolerance-related  
a crop in a field -

PS Claim 13; Fig 4; 92pp; English.

CC The present sequence represents a plant stress tolerance-related  
myloblastosis (MYB) transcription factor. The specification describes  
AtMYB60, AtMYB74, AtMYB75 and AtMYB90. MYB polypeptides are useful for  
enhancing a plant's tolerance or sensitivity to stress e.g. salt,  
drought, cold and heat tolerance. It is also useful for producing a  
transgenic plant with enhanced stress tolerance or stress sensitivity.

CC MYB polypeptides are also useful for increasing the production of  
CC products of the phenylpropanoid biosynthesis pathway, e.g. stilbenes,  
CC flavonoids, lignins, salicylic acid, anthocyanins, and phenolic  
CC derivatives.

XX Sequence 249 AA;

Query Match 36.2%; Score 535; DB 22; Length 249;  
Best Local Similarity 47.2%; Pred. No. 5.6e-43;  
Matches 117; Conservative 27; Mismatches 64; Indels 40; Gaps 6;

QY 7 SSGVKGKSWTDEEDFLRKCIDKYGKXHLVPIRAGLNCRKSCRLRLNTLRLPHIKR 66  
DB 4 SSKGLRGKAWTDEEDSLRQICINKYEGKXHVPLRAGLNCRKSCRLRLNTLRLPHIKR 63  
QY 67 GDFODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLRLKNTTIVPREKI 126  
DB 64 GKLSNDEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLSSKHSCKSKMKK 123  
QY 127 NNKGEIST--KIEIIPQRRKY-----FSTWKVNTNNVIL 162  
DB 124 KNIISPPTPVQKIGVFKPRPSFTVNNDCNHLNAPRYDVNP-----CLGLINNNVCDNS 179  
QY 163 DEEHECKEIISEKQTPDASNDVDPWMT-NL-ENCNDIIEDEEVYINYEKTLT----- 215  
DB 184 NKDEXKDDFVN-----NLMGNDNMWLENLAGEHDAIVEATTAEGATLADVEQ 236  
QY 216 --SLTHEE 221  
DB 237 LMSLFHGE 244

RESULT 6  
AAB01916  
ID AAB01916 standard; Protein; 249 AA.

XX AAB01916;

DT 31-JUL-2001 (first entry)

DE Arabidopsis thaliana transcription factor G663.

KM Transcription factor; trait modification; seed characteristic;  
KW structural characteristic; developmental characteristic; gene therapy;  
KM agricultural biotechnology.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers

FT Domain 9..111 /label=Conserved\_domain

XX MO200135727-A1.

XX 25-MAY-2001.

XX 14-NOV-2000; 2000WO-US31457.

XX 17-NOV-1999; 99US-0166228.

XX 17-APR-2000; 2000US-0197899.

XX 22-AUG-2000; 2000US-0227439.

XX (MENDEL) MENDEL BIOTECHNOLOGY INC.

XX (REUBER) REUBER L.

XX (CREE) CREEBMAN R.

XX (PLIG/) PILGRIM M.

XX (RIEC/) RIECHMANN J L.

XX (JIANG/) JIANG C.

XX (YUGG/) YU G.

XX (PINE/) PINEDA O.

XX (HEAR/) HEARD J.

XX Reuber L, Creelman R, Pilgrim M, Riechmann J, Jiang C, Yu G;





Oy	127	NNKGEISF---KIEIKQRKRKFSSTMNKNTNNVLLDEERHCKEIIISKQTPDASMD	188
Db	124	KNIIISPPTPVOKIGVFKRPRSF---SVNNGCSHLNGLPBYDLIRSCGLKKN-----	174
Oy	184	NVDPWMINLEN---CNDIEEDEEVIVYKTLTSLHREISPPINIGSGNSMOOQIS	240
Db	175	-----INVCENSITCNKDEKDDP-----NNLNNGDMWMLLENLL	208
Oy	241	HENWGEPFSLNPPMOOGVONDDPSAID-LNNLID	274
Db	209	GEN-OEADVAIVEATTAEHGATLAFPEVLOLWSIFD	242
RESULT 8			
AAE30046			
ID	AAE30046	standard; Protein; 249 AA.	
AC	AAE30046;		
DT	24-FEB-2003	(first entry)	
DE	Arabidopsis thaliana G663	transcription factor protein.	
KW	Transcription factor; metabolite pathway; terpenoid; limonene synthase;		
KW	alkaloid pathway gene; taxadiene synthase; biological pathway; freezing;		
KW	abiotic stress; cold; drought; heat; nutrient deficiency; biotic stress;		
KW	infection; developmental pathway; flowering; root development; TDS; LS;		
KW	transgenic; transgenic plant.		
OS	Arabidopsis thaliana.		
PN	WO200274917-A2.		
PD	26-SEP-2002.		
PF	15-MAR-2002; 2002MO-US07999.		
PR	16-MAR-2001; 2001US-0810836.		
PA	(MEND-) MENDEL BIOTECHNOLOGY INC.		
PI	Brown P;		
DR	WPI, 2003-018799/01.		
XX	N-PSDB; AAD47497.		
PT	Determining whether one of several test transcription factor (TF)		
PT	polynucleotides encodes pathway TF by determining expression from		
PT	pathway gene promoter linked to reporter gene in a cell in presence of		
PT	test polynucleotides -		
PS	Disclosure; Page 153; 221pp; English.		
XX	The present invention relates to a high-throughput method for identifying		
CC	a polynucleotide which encodes a transcription factor for controlling the		
CC	expression of one or more genes in a pathway. The method is useful for		
CC	determining whether a member of a pool of test transcription factor		
CC	polynucleotides encodes a biosynthetic pathway transcription factor or		
CC	a primary metabolite pathway gene or a secondary metabolite pathway gene		
CC	such as a terpenoid or alkaloid pathway gene transcription factor. It		
CC	is also useful for determining whether a member of a pool of test encodes		
CC	a mentha sp. or taxus sp. terpenoid pathway gene where the terpenoid		
CC	pathway gene encodes limonene synthase (LS) or taxadiene synthase (TDS).		
CC	It is also useful for identifying one or more transcription factors that		
CC	activate one or more genes of a biological pathway of a plant, fungi or		
CC	animal cell, where the biological pathway can be a biochemical pathway		
CC	such as biosynthetic pathways for soluble and insoluble carbohydrates,		
CC	for amino acids and a response pathway to abiotic stress such as cold,		
CC	freezing, drought, heat, nutrient deficiency, pH or biotic stress such		
CC	as fungal, viral or bacterial infection, developmental pathway such as		
CC	flowering, root development, a response pathway to environmental cues		
CC	such as light intensity and light quality, circadian rhythm. Sequences		
CC	of the invention are used to generate transgenic plants. The present		

[illegible]

ID	AAU75735	standard; Protein; 249 AA.
AC	AAU75735;	
DT	08-MAY-2002	(first entry)
DE	A. thaliana Production of anthocyanin pigment 2 (PAP2) protein.	
XX		
XX	PAP2; production of anthocyanin pigment; MYB-like transcription factor;	
KM	transgenic; plant; phenylpropanoid; stress; light stress; water stress;	
KM	pH stress; temperature stress; heavy metal stress; pathogen attack;	
KM	infection; wounding; nutrient deficiency; herbivory; plant colouration;	
XX	stress sensitisation; chromosome 1.	
OS	Arabidopsis thaliana.	
XX		
FM	Key	location/Qualifiers
FT	misc_difference 130	/note= "Encoded by CCT"
FT	misc_difference 161	/note= "Encoded by GAA"
FT	misc_difference 190	/note= "Encoded by GAT"
FT	misc_difference 215	/note= "Encoded by GAT"
FT	misc_difference 244	/note= "Encoded by GAT"
FT	misc_difference 244	/note= "Encoded by GAG"
XX		
PN	WO200200902-A2.	
XX		
PD	03-JAN-2002.	
XX		
PF	21-JUN-2001;	2001WO-US19734.
XX		
XX	23-JUN-2000;	2000US-0603244.
PR	05-JUL-2000;	2000US-0610185.
XX		
XX	(SALK ) SALK INST BIOLOGICAL STUDIES.	
PA	(ROBE-) ROBERTS NOBLE FOUND INC SAMUEL.	
XX		
PI	Borevitz J,	Xia Y, Lamb CJ, Dixon RA;
XX		

DR WPI, 2002-164443/21.  
 DR N-PSDB; ABK14222.  
 PT New isolated PAP1 or PAP2 gene, useful for increasing pigmentation in  
 PT plants, as reporter genes for analysing expression pattern of promoter  
 PT of interest, and to increase flux through phenylpropanoid pathway  
 PS Claim 9; Page 28; 29pp; English.  
 XX  
 CC This invention relates to the nucleotide and protein sequences of novel  
 CC Arabidopsis thaliana production of anthocyanin pigment (PAP) genes PAP1  
 CC or PAP2. The genes for both PAP1 and PAP2 have been mapped to chromosome  
 CC 1. PAP1 and PAP2 are MYB-like transcription factors that regulate the  
 CC production of anthocyanin pigment. The nucleotide sequence of the  
 CC invention is useful for screening for PAP gene expression, it is also  
 CC useful for detecting promoter activity within a plant cell. The  
 CC nucleotide sequence is also useful for screening for plant cell (e.g.,  
 CC A. thaliana) transformation. The transformed plant cells are then used  
 CC to produce transgenic plants with tissue-specific altered colour  
 CC expression. The nucleotide sequence is also useful for enhancing  
 CC accumulation of phenylpropanoid product in a plant. A transgenic  
 CC plant is useful for detecting a stress condition such as light  
 CC stress, water stress, pH stress, temperature stress, heavy metal  
 CC stress, pathogen attack or infection, wounding, nutrient deficiency,  
 CC herbivory, or abnormal hormone levels in a plant. PAP1 or PAP2  
 CC overexpression can also be useful in developing or enhancing desirable  
 CC colouration in ornamental plants for commercial sales, and for  
 CC sensistising plants to stress. Introduction of PAP1 or PAP2 sensitises  
 CC plants to respond more quickly to a stress condition in ways that are  
 CC physiologically significant, as well as in ways of ornamental  
 CC significance. The present sequence represents the A. thaliana PAP2  
 CC protein sequence of the invention.  
 CC  
 SQ Sequence 249 AA;

Query Match 35.7%; Score 527; DB 23; Length 249;  
 Best Local Similarity 48.4%; Pred. No. 3.3e-42;  
 Matches 109; Conservative 28; Mismatches 54; Indels 34; Gaps 4;

QY 7 SSIGVKGSTWDEEDFLRKICIDYKGGKWHLYPIRAGLNCRKSCRLRWNTLRPHIKR 66  
 DB 4 SSKGLRGAWTAEEDSLRLCICIDYKGGKWHLYPIRAGLNCRKSCRLRWNTLRPHIKR 63  
 QY 67 GDFEODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKYWNNTLRLKTLTKIVPEKI 126  
 DB 64 GRUSNDEVLDLRLHKLGNRWSLIAGRLPGRTANDVKYWNNTLRLKTLTKIVPEKI 123  
 QY 127 NNKCGEIST--KIIIKPORKY-----FSTYKQNTVNNVIL 162  
 DB 124 KNISPTTVOXKIGVPRRSVANGCSHLNGLPXVDLIPSCGLAKKNVENSITC 183  
 QY 163 DEEBHCKEISEKOTPDASMDNDVPMWNLINLNCNDIDEEDEV 207  
 DB 184 NKDEKXDFVN-----NLMGNDMW---LENILGENOXEAXLV 218

RESULT 10  
 AAU93155 standard; Protein; 212 AA.  
 AAU93155;  
 02-JUL-2002 (first entry)  
 Arabidopsis transcription factor #193.  
 Agricultural; metabolic chemical; environmental stress; drought;  
 microbial disease resistance; herbicide resistance; seed yield;  
 fruit yield; growth rate; leaf senescence; flower senescence.  
 plant; transcription factor; transgenic.  
 Arabidopsis thaliana.

PN WO200215675-A1.  
 XX 28-FEB-2002.  
 PD 22-AUG-2001; 2001WO-US26189.  
 XX 22-AUG-2000; 2000US-227439P.  
 XX 16-NOV-2000; 2000US-0713994.  
 PR 16-APR-2001; 2001US-0837944.  
 XX  
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (PIIG/) PILGRIM M.  
 PA (CREE/) CREELMAN R.  
 PA (DUBE/) DUBELL A J.  
 PA (HEAR/) HEARD J L.  
 PA (JIAN/) JIANG C.  
 PA (KEDJ/) KEDDIE J.  
 PA (ADAM/) ADAM L.  
 PA (RATC/) RATCLIFF O.  
 PA (REUB/) REUBER J L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (YUGG/) YU G.  
 PA (PINE/) PINEDA O.  
 XX  
 PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J,  
 PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
 DR WPI, 2002-292022/33.  
 DR N-PSDB; ABK65341.  
 XX  
 PT An isolated or recombinant polynucleotide used to produce a transgenic  
 PT plant -  
 XX  
 SQ Claim 40; Page 794-795; 941pp; English.

CC The invention relates to 1 of 232 isolated or recombinant polynucleotides  
 CC encoding an Arabidopsis thaliana transcription factor, their variants,  
 CC complements, fragments, or related polynucleotide with 31% to 95%  
 CC sequence identity, where the plant possesses an altered trait as compared  
 CC to a wild-type or reference plant, or the plant exhibits an altered  
 CC phenotype as compared to a wild-type or reference plant, or the plant  
 CC exhibits ectopic expression or altered expression of one or more genes  
 CC associated with a plant trait as compared to a wild plant. Also included  
 CC are a transgenic plant comprising the polynucleotides, a computer  
 CC readable medium having stored sequence information, and identifying a  
 CC homologue sequence from a database comprising a plurality of known plant  
 CC sequences comprising inputting sequence information selected from one of  
 CC 464 fully defined sequences given in the specification. The isolated or  
 CC recombinant polynucleotide is used for producing a plant having a  
 CC modified trait, the method comprising selecting a polynucleotide that  
 CC encodes a polypeptide or an antisense nucleic acid, inserting the  
 CC polynucleotide or antisense nucleic acid into an expression vector,  
 CC introducing the vector into a plant or a cell of a plant to overexpress  
 CC the polypeptide or antisense nucleic acid, thereby producing a modified  
 CC plant, and selecting for a modified trait (e.g. increased  
 CC production of agriculturally useful proteins or metabolic chemicals,  
 CC pest tolerance, environmental stress response (e.g. drought), microbial  
 CC disease resistance, herbicide resistance, seed and fruit yield, growth  
 CC rate, leaf and flower senescence and many other traits listed in the  
 CC specification). The present sequence is one of the 232 proteins which are  
 CC A. thaliana transcription factors.  
 CC  
 XX

Query Match 35.5%; Score 524; DB 23; Length 212;  
 Best Local Similarity 48.3%; Pred. No. 5e-42;  
 Matches 102; Conservative 33; Mismatches 38; Indels 38; Gaps 5;

QY 7 SSIGVKGSTWDEEDFLRKICIDYKGGKWHLYPIRAGLNCRKSCRLRWNTLRPHIKR 66  
 DB 4 SSKGLRGAWTAEEDSLRLCICIDYKGGKWHLYPIRAGLNCRKSCRLRWNTLRPHIKR 63  
 QY 67 GDFEODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKYWNNTLRLKTLTKIVPEKI 114

```

Db      64  GKLSDEVDLLRLHRLGNRWSLIAGRLPGRTANDVKNYWNTHLSKHEPCKIKMKKR 123
Qy      115  -----LNTKTIYPREK---INNKGEIST--RIEIIKPKQRKCFSTIMKNVNNN 159
Db      124  DITPIPTPALKNVNYKPRPSFTVNDCHLNAPKVDVNP---CLGLTINNVCDS 179
Qy      160  VILDEEHCKEIIIEKOTPDASMDNVPMWI 190
Db      180  IIVYKDKKKQDLV-----NLIDGNMML 203

RESULT 11
AAE01933
ID  AAE01933 standard; Protein, 209 AA.
XX
AC  AAE01933;
XX
DT  31-JUL-2001 (first entry)
XX
DE  Arabidopsis thaliana transcription factor homologue G2421.
XX
KM  Transcription factor; trait modification; seed characteristic;
KM  structural characteristic; developmental characteristic; gene therapy;
KM  agricultural biotechnology.
XX
OS  Arabidopsis thaliana.
XX
FH  Key
FT  Domain
FT  Location/Qualifiers
XX  9..110
XX  /label=Conserved_domain

WO200135727-A1.
XX
PD  25-MAY-2001.
XX
PF  14-NOV-2000; 2000WO-US31457.
XX
PR  17-NOV-1999; 99US-0166228.
PR  17-APR-2000; 2000US-0187889.
PR  22-AUG-2000; 2000US-0227439.
XX
PA  (MEND-) MENDEL BIOTECHNOLOGY INC.
PA  (REUB/) REUBER L.
PA  (CREE/) CREELMAN R.
PA  (PILG/) PILGRIM M.
PA  (RIEC/) RIECHMANN J L.
PA  (JIANG/) JIANG C.
PA  (YUGG/) YU G.
PA  (PINE/) PINEDA O.
PA  (HEAR/) HEARD J.
XX
PI  Reuber L, Creelman R, Pilgrim M, Riechmann JL, Jiang C, Yu G,
PI  Pineda O, Heard J;
XX
DR  WPI; 2001-335979/35.
DR  N-RSDB; AAD05768.
XX
PT  Nucleic acid encoding plant transcription factor polypeptides, useful
PT  for altering the developmental and structural characteristics of
PT  plants, e.g. corn, potato and cotton plants -
XX
PS  Claim 4; Page 116; 133pp; English.
XX
CC  The present sequence is homologue G2421 of Arabidopsis thaliana
CC  transcription factor G663. The transcription factors may be used to
CC  modify traits associated with structural or developmental
CC  characteristics of plants, e.g., corn, potato and cotton plants, when
CC  their expression level is altered. Specifically, they are used for
CC  modifying the nutritional content of plants or seeds, e.g., to modify
CC  vitamin and mineral content, to modify the oil content of seeds, to
CC  modify the insoluble sugar content of seeds, to modify phenyl lipid
CC  content, to modify fatty acid and modify wax content. They may also be

```

```

CC  used to alter seed characteristics such as shelf-life, size, stress
CC  tolerance, seedling vigour, pest and pathogen resistance and germination
CC  rate. They may also be used in gene therapy. Therefore manipulating
CC  transcription factor levels in plants offers great potential in
CC  agricultural biotechnology for modifying a plant's traits.
XX
SQ  Sequence 209 AA;
XX
Query Match 35.1%; Score 518.5; DB 22; Length 209;
Best Local Similarity 44.0%; Pred. No. 1,7e-41;
Matches 120; Conservative 29; Mismatches 45; Indels 79; Gaps 11;

Qy      7  SSLGVRKSWTDEDFLLRKCIDKYGEGKXHLVPIRAGLNRCRCKSLRLMNTLPHIKR 66
Db      4  SSKGRKGAWTAEDSLRQCICKYGEKGKHQVPLBAGLNRCRCKSLRLMNTLPHIKR 63
Qy      67  GDFDEVDLLRLHRLGNRWSLIAGRLPGRTANDVKNYWNTHLSKHEPCKIKMKKR 126
Db      64  GKFSDEVDLLRLHRLGNRWSLIAGRLPGRTANDVKNYWNTHLSKHEPCKIKMKKR 111
Qy      127  NNRKGEISTK-IEIIKPKQRKCFSTIMKNVNNVILDEEHCKEIIIEKOTPDASMDNV 185
Db      112  HECCCKTKIRIITPP-----NTPAQKVCESITCNKDD-----EK---DIPVDNF 156
Qy      186  ---DPWMINLLENCNDIEDEEV-VINYEKVLTSLHEISPLNIGSGNSMQGQISH 241
Db      157  MVGDNIWLRL-----LDGQGVDLVTEAATE-----KEGTIA- 191
Qy      242  ENNGEFSNLNPMQGVQNDPFAEIDVNLWD 274
Db      192  -----FDVE-QLWNLFD 202

RESULT 12
AAU93176
ID  AAU93176 standard; Protein, 209 AA.
XX
AC  AAU93176;
XX
DT  02-JUL-2002 (first entry)
XX
DE  Arabidopsis transcription factor #214.
XX
KM  Agriculture; metabolic chemical; environmental stress; drought;
KM  microbial disease resistance; herbicide resistance; seed yield;
KM  fruit yield; growth rate; leaf senescence; flower senescence.
XX
OS  Arabidopsis thaliana.
XX
PN  WO200215675-A1.
XX
PD  28-FEB-2002.
XX
PF  22-AUG-2001; 2001WO-US26189.
XX
PR  22-AUG-2000; 2000US-227439P.
PR  16-NOV-2000; 2000US-0713944.
PR  16-APR-2001; 2001US-0837944.
XX
PA  (MEND-) MENDEL BIOTECHNOLOGY INC.
PA  (PILG/) PILGRIM M.
PA  (CREE/) CREELMAN R.
PA  (DUBE/) DUBELL A J.
PA  (HEAR/) HEARD J.
PA  (JIANG/) JIANG C.
PA  (KEDD/) KEDDIE J.
PA  (ADAM/) ADAM L.
PA  (RATC/) RATCLIFF O.
PA  (REUB/) REUBER J L.
PA  (RIEC/) RIECHMANN J L.
PA  (YUGG/) YU G.
PA  (PINE/) PINEDA O.

```

XX pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;  
 PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
 XX WPI, 2002-292022/33.  
 DR N-PSDB; ABK65362.  
 XX An isolated or recombinant polynucleotide used to produce a transgenic  
 PT plant -  
 XX  
 PS Claim 40; Page 878-879; 941pp; English.  
 CC The invention relates to 1 of 232 isolated or recombinant polynucleotides  
 CC encoding an Arabidopsis thaliana transcription factor, their variants,  
 CC complements, fragments, or related polynucleotide with 31% to 95%  
 CC sequence identity, where the plant possesses an altered trait as compared  
 CC to a wild-type or reference plant, or the plant exhibits an altered  
 CC phenotype as compared to a wild-type or reference plant, or the plant  
 CC exhibits ectopic expression or altered expression of one or more genes  
 CC associated with a plant trait as compared to a wild plant. Also included  
 CC are a transgenic plant comprising the polynucleotide, a computer  
 CC readable medium having stored sequence information, and identifying a  
 CC homologue sequence from a database comprising a plurality of known plant  
 CC sequences comprising inputting sequence information selected from one of  
 CC 464 fully defined sequences given in the specification. The isolated or  
 CC recombinant polynucleotide is used for producing a plant having a  
 CC modified trait, the method comprising selecting a polynucleotide that  
 CC encodes a polypeptide or an antisense nucleic acid, inserting the  
 CC polynucleotide or antisense nucleic acid into an expression vector,  
 CC introducing the vector into a plant or a cell of a plant to overexpress  
 CC the polypeptide or antisense nucleic acid, thereby producing a modified  
 CC plant, and selecting for a modified trait (e.g. increased  
 CC production of agriculturally useful proteins or metabolic chemicals,  
 CC pest tolerance, environmental stress response (e.g. drought), microbial  
 CC disease resistance, herbicide resistance, seed and fruit yield, growth  
 CC rate, leaf and flower senescence and many other traits listed in the  
 CC specification). The present sequence is one of the 232 proteins which are  
 CC A. thaliana transcription factors.  
 CC  
 XX  
 SQ Sequence 209 AA;  
 Query Match 35.1%; Score 518.5; DB 23; Length 209;  
 Best Local Similarity 44.0%; Pred. No. 1,7e-41;  
 Matches 120; Conservative 29; Mismatches 45; Indels 79; Gaps 11;  
 QY 7 SSGVKGKSWTDEEDFLKRCIDKYGKWHLYPIRAGLNRCRCKSLRWLNTLRPHIKR 66  
 DB 4 SSKGLRKGAWTAEEDSLRQCIGYGGKWHQVPLRAGLNRCRCKSLRWLNTLRPHIKR 63  
 QY 67 GDEODEVDLILRLHKLGNRWSLIGRLPGRTANDVKNYWNNTLRKNTTKIVPREKI 126  
 DB 64 GKSSDSVDLILRLHKLGNRWSLIGRLPGRTANDVKNYWNNTLRKNTTKIVPREKI 111  
 QY 127 NNKCGEISTK-ILIKPQRKRYFSSTMKNTNNVILDEEHCKEIISEKQTPDASMDNV 185  
 DB 112 HEPCCKTKIKRINIITPP-----NTPAQKXCENSITCKMD-----EK---DDPFDNF 156  
 QY 186 ---DPMWNLLENQNDIEDEEV-VINYEKTLTSLIHEBISPLNIGNSNQOQISH 241  
 DB 157 MVDGNLWLERL-----LDEGQEVDLVTEAATE-----KEGTIA- 191  
 QY 242 ENWGEPSLNPPOGQGVNDPFAEIDLMKLD 274  
 DB 192 -----FDVE-OLMNLFD 202  
 RESULT 13  
 ID AAB68355 standard; Protein; 211 AA.  
 AC AAB68355;  
 XX 09-JUL-2001 (first entry)

XX Amino acid sequence of MYB transcription factor AtMYB75.  
 DE Stress tolerance; myb1a stress transcription factor; heat tolerance;  
 XX MYB transcription factor; AtMYB60; AtMYB74; AtMYB75; transgenic plant;  
 KW phenylpropanoid biosynthesis.  
 XX Arabidopsis sp.  
 XX  
 OS Key  
 FT Misc-difference 72 Location/Qualifiers  
 FT Misc-difference 72 /note= "asp encoded by TTTCGAT"  
 FT Misc-difference 90 /note= "gly encoded by TTTCGA"  
 FT Misc-difference 173 /note= "asn encoded by TAAAT"  
 FT Misc-difference 180 /note= "his encoded by ATC"  
 FT Misc-difference 181 /note= "tyr encoded by ATATAC"  
 XX  
 XX NO200132002-A1.  
 XX  
 PD 10-MAY-2001.  
 XX  
 PF 06-NOV-2000; 2000MO-US30503.  
 XX  
 PR 05-NOV-1999; 99US-0163579.  
 PR 23-OCT-2000; 2000US-0693855.  
 XX  
 PA (BADI ) BASF CORP.  
 XX  
 PI Tonelli C;  
 XX  
 XX WPI, 2001-316365/33.  
 DR N-PSDB; AAF85190.  
 XX  
 PT Nucleic acid molecules encoding plant stress tolerance-related  
 PT myb1a stress transcription factors for increasing stress resistance of  
 PT a crop in a field -  
 XX  
 PS Claim 13; Fig 3B; 92pp; English.  
 XX  
 CC The present sequence represents a plant stress tolerance-related  
 CC myb1a stress transcription factor. The specification describes  
 CC AtMYB60, AtMYB74, AtMYB75 and AtMYB90. MYB polypeptides are useful for  
 CC enhancing a plant's tolerance or sensitivity to stress e.g. "salt",  
 CC drought, cold and heat tolerance. It is also useful for producing a  
 CC transgenic plant with enhanced stress tolerance or stress sensitivity.  
 CC MYB polypeptides are also useful for increasing the production of  
 CC products of the phenylpropanoid biosynthesis pathway, e.g. stilbenes,  
 CC flavonoids, lignins, salicylic acid, anthocyanins, and phenolic  
 CC derivatives.  
 CC  
 XX  
 SQ Sequence 211 AA;  
 Query Match 34.5%; Score 510; DB 22; Length 211;  
 Best Local Similarity 49.3%; Pred. No. 1.1e-40;  
 Matches 103; Conservative 25; Mismatches 43; Indels 38; Gaps 5;  
 QY 7 SSGVKGKSWTDEEDFLKRCIDKYGKWHLYPIRAGLNRCRCKSLRWLNTLRPHIKR 66  
 DB 4 SSKGLRKGAWTAEEDSLRQCIGYGGKWHQVPLRAGLNRCRCKSLRWLNTLRPHIKR 63  
 QY 67 GDEODEVDLILRLHKLGNRWSLIGRLPGRTANDVKNYWNNTLRKNTTKIVPREKI 114  
 DB 64 GKSSDSVDLILRLHKLGNRWSLIGRLPGRTANDVKNYWNNTLRKNTTKIVPREKI 123  
 QY 115 -----LNTTKIVPREK---INNKCGEIST-KIILKQRKRYFSSTMKNTNN 159  
 DB 124 DITPIPTPALKNVYPPRPSFTVNNDCNHLNAPKVVNP-----CLGLNINNVCDNS 179  
 QY 160 VILDEEHCKEIISEKQTPDASMDNVDPW 168



```

Qy 124 ---EKINNKGEISTXIEIKPORRKYFSSTMKNVTNNVILDEEHCKEIISEKOTPD 180
Db 130 SESEDICSSPG---NSEVSRKSGRENNAEIPRKVADGAVDIDKE---EDITEDQTSAG 182
Qy 181 SMDNVDPWWINLJENCN-----DIEDEEV---VINEKTLTSLHHEISPLINIG 229
Db 183 LPEN-----QLETSNSQCPSVATDFVPOAPSIPTAYSFQOSTTSSVPGVSDVDVN 236
Qy 230 EGNMGOQGISHENWGEFSLNLPMOGVQND 261
Db 237 HNKGSKQVPPPLSTNACFNSS---AQGVAGD 264

```

Search completed: January 29, 2004, 20:11:24  
 Job time : 43 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 20:10:56 ; Search time 21 Seconds  
(without alignments)  
552.056 Million cell updates/sec

Title: US-10-033-190-2  
Perfect score: 1477  
Sequence: 1 NMSTMSLSGVKSGWTDEE.....QQGVNDPFAEIDWNLTD 274

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues  
Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Issued Patents AA: \*  
2: /cgn2\_6/prodata/1/aa/5A COMB.pep: \*  
3: /cgn2\_6/prodata/1/aa/5B COMB.pep: \*  
4: /cgn2\_6/prodata/1/aa/6A COMB.pep: \*  
5: /cgn2\_6/prodata/1/aa/6B COMB.pep: \*  
6: /cgn2\_6/prodata/1/aa/6C COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	538	36.4	248	4	US-09-610-185C-2	Sequence 2, Appl1
2	527	35.7	249	4	US-09-610-185C-4	Sequence 4, Appl1
3	370.5	25.1	278	2	US-08-722-626B-2	Sequence 2, Appl1
4	317.5	21.5	553	3	US-08-997-251-2	Sequence 2, Appl1
5	315	21.3	553	3	US-08-997-251-4	Sequence 4, Appl1
6	242.5	16.4	751	4	US-09-402-929-2	Sequence 2, Appl1
7	242.5	16.4	752	4	US-09-402-929-6	Sequence 4, Appl1
8	240.5	16.3	156	3	US-08-928-941D-4	Sequence 6, Appl1
9	240.5	16.3	156	3	US-09-280-590A-4	Sequence 4, Appl1
10	236.5	16.0	156	3	US-08-928-941D-36	Sequence 36, Appl1
11	236.5	16.0	156	3	US-09-280-590A-46	Sequence 46, Appl1
12	157.5	10.7	802	3	US-09-156-316-1	Sequence 1, Appl1
13	138	9.3	51	1	US-08-519-103-16	Sequence 16, Appl1
14	138	9.3	51	1	US-08-519-103-17	Sequence 17, Appl1
15	138	9.3	51	1	US-09-018-635-16	Sequence 16, Appl1
16	138	9.3	51	3	US-09-018-635-17	Sequence 17, Appl1
17	138	9.3	51	4	US-09-018-635-15	Sequence 15, Appl1
18	138	9.3	51	4	US-09-912-962-17	Sequence 17, Appl1
19	135	9.1	51	4	US-09-912-962-16	Sequence 16, Appl1
20	135	9.1	51	1	US-08-519-103-18	Sequence 18, Appl1
21	135	9.1	51	3	US-09-018-635-18	Sequence 18, Appl1
22	132	8.9	94	2	US-09-912-962-18	Sequence 18, Appl1
23	132	8.9	94	4	US-08-814-030-1	Sequence 1, Appl1
24	131	8.9	51	1	US-08-519-103-19	Sequence 19, Appl1
25	131	8.9	51	3	US-09-018-635-19	Sequence 19, Appl1
26	127.5	8.6	372	3	US-09-912-962-19	Sequence 19, Appl1
27	127.5	8.6	372	3	US-08-928-941D-16	Sequence 16, Appl1

28	127.5	8.6	761	3	US-08-928-941D-1	Sequence 1, Appl1
29	127.5	8.6	761	4	US-09-280-590A-1	Sequence 1, Appl1
30	119	8.1	760	4	US-08-928-941D-29	Sequence 29, Appl1
31	119	8.1	760	4	US-09-280-590A-29	Sequence 29, Appl1
32	109.5	7.4	169	3	US-08-928-941D-35	Sequence 35, Appl1
33	109.5	7.4	169	4	US-09-280-590A-45	Sequence 45, Appl1
34	104.5	7.1	52	1	US-08-519-103-14	Sequence 14, Appl1
35	104.5	7.1	52	4	US-09-018-635-14	Sequence 14, Appl1
36	104.5	7.1	52	4	US-09-912-962-14	Sequence 14, Appl1
37	103.5	7.0	52	1	US-09-156-316-5	Sequence 5, Appl1
38	103.5	7.0	52	1	US-08-519-103-13	Sequence 13, Appl1
39	103.5	7.0	52	3	US-09-018-635-13	Sequence 13, Appl1
40	103.5	7.0	52	4	US-09-912-962-13	Sequence 13, Appl1
41	100.5	6.8	50	3	US-09-156-316-4	Sequence 4, Appl1
42	100.5	6.8	50	3	US-09-156-316-6	Sequence 6, Appl1
43	94	6.4	507	4	US-09-153-277-4	Sequence 4, Appl1
44	94	6.4	695	4	US-09-153-277-2	Sequence 4, Appl1
45	92.5	6.3	1664	1	US-09-599-652-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1  
US-09-610-185C-2  
; Sequence 2, Application US/09610185C  
; Patent No. 6573432  
; GENERAL INFORMATION:  
; APPLICANT: Borevitz, Uetlin  
; APPLICANT: Xia, Yijl  
; APPLICANT: Dixon, Richard A.  
; APPLICANT: Lamb, Christopher J.  
; TITLE OF INVENTION: REGULATION OF ANTHOCYANIN PIGMENT  
; FILE REFERENCE: SALKINS: 003C1  
; CURRENT APPLICATION NUMBER: US/09/610,185C  
; CURRENT FILING DATE: 2000-07-05  
; PRIOR APPLICATION NUMBER: US 09/603,244  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 248  
; TYPE: PRT  
; ORGANISM: Arabidopsis Thaliana  
US-09-610-185C-2

Query Match 36.4%; Score 538; DB 4; Length 248;  
Best Local Similarity 38.9%; Pred. No. 2,2e+48;  
Matches 115; Conservative 47; Mismatches 48; Indels 86; Gaps 10;  
QY 7 SSGVGRKSGWTDEEDFLKCKIDKYGEGKWLVPVPAAGLNCRKSKRLMNTLPHIKR 66  
DB 4 SSGKLGKGMATTEEDDLKQCKINIKYGEKWHQVPAAGLNCRKSKRLMNTLPHIKR 63  
QY 67 GDEODEVLLIRHLGLNRMWSLNGRLPGFRANVKNYNNL-----LRKL 115  
DB 64 GKLSSEVDLRLHRLGLNRMWSLNGRLPGFRANVKNYNNLHLSKKEPPCKIKKI 123  
QY 116 NTKI-----VPREK--INNKGEIST--KIELIKPQRKRYFSSTKNTNNN 159  
DB 124 DTRPITPAKNNVYKPRPRSTVANNDCNHLNAPKVVNP-----CGLNINNVCDNS 179  
QY 160 VILDEEHCKEITISEKQTPDASHMDVDPWVWINILNENDIEDEEVVINEYKTLTSLH 219  
DB 180 IYNNKKKKDQLVN-----MLIDGDNWLEKF-----LBSQV----- 213  
QY 220 EETSPPLTIGEGSMQGGQISHENWGEFSLNLPWQGVNDPFAEID-LMWLTD 274  
DB 214 -----DILVPEATTEKG-----DTLARDVDQLMSLFD 241  
RESULT 2

```

US-09-610-185C-4
Sequence 4, Application US/09610185C
Patent No. 6573432
GENERAL INFORMATION:
APPLICANT: Borevitz, Justin
APPLICANT: Xia, Yiji
APPLICANT: Dixon, Richard A.
APPLICANT: Lamb, Christopher J.
TITLE OF INVENTION: REGULATION OF ANTHOCYANIN PIGMENT
FILE REFERENCE: SALKINS 003C1
CURRENT APPLICATION NUMBER: US/09/610,185C
CURRENT FILING DATE: 2000-07-05
PRIOR APPLICATION NUMBER: US 09/603,244
PRIOR FILING DATE: 2000-06-23
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 249
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURES:
NAME/KEY: VARIANT
LOCATION: (1)...(249)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-610-185C-4

```

Query Match	35.7%;	Score 527;	DB 4;	Length 249;
Best Local Similarity	48.4%;	Pred. No. 3.1e-47;		
Matches 109;	Conservative 28;	Mismatches 54;	Indels 34;	Gaps 4;

```

QY 7 SLSLGVKGSWTDEDEFLRKCIDCKGEGSMHVLPRAGLNCRCSCRJMYLNYLRPHIKR 66
D 7 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 4 SSKGLRKGAATTAEDBSLRLCIDCKYGEGRHGVPLRAGLNCRCSCRJMYLNYLRPHIKR 63
QY 67 GDFEQDEVDLLRLRLKLLGNRWSLAGRLPGRTANDVKNYNNYLRLKRLNTTKTVPREKI 126
D 64 GRLSNDEVDLLRLRLKLLGNRWSLAGRLPGRTANDVKNYNNYLRLSKHGESSCKSKHKK 123
QY 127 MNKCGEIST---KLEILKQRRKY-----FSTMTQNNNNNVIL 162
D 124 KNISPTTFOVKITGVFKPRPRFSFVNNGCSHLNGLPYVDLPSGLGLKNNVCENSITC 183
QY 163 DEEBHCKEILSEKQTPDASMDNVDPWMLLNCNDIDIEDEEYV 207
D 184 NKDEKKDFVN-----NLNMGDNNM---LENLIGNEDEAKATV 218

```

RESULT 3  
 US-08-722-626B-2  
 Sequence 2, Application US/08722626B  
 Patent No. 5939601  
 GENERAL INFORMATION:  
 APPLICANT: Yang, Yimong  
 APPLICANT: Kleszig, Daniel, F.  
 TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ENHANCED  
 TITLE OF INVENTION: DISEASE RESISTANCE IN PLANTS  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Dann, Dorfman, Herrell and Skillman  
 STREET: 1601 Market Street Suite 720  
 CITY: Philadelphia  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19103-2307  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/722,626B  
 FILING DATE: 27-SEP-1996

```

1 CLASSIFICATION: 800
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER:
4 FILING DATE:
5 ATTORNEY/AGENT INFORMATION:
6 NAME: Pat Hagan
7 REGISTRATION NUMBER: 27,643
8 REFERENCE/DOCKET NUMBER: 97-0010
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: 215 563-4100
11 TELEFAX: 215 563-4044
12
13 TELERX:
14
15 INFORMATION FOR SEQ ID NO: 2:
16
17 SEQUENCE CHARACTERISTICS:
18 LENGTH: 278 amino acids
19 TYPE: amino acid
20 STRANDEDNESS: single
21 TOPOLOGY: linear
22 MOLECULE TYPE: protein
23 HYPOTHEetical: NO
24 ANTI-SENSE: NO
25 FRAGMENT TYPE: internal
26
27 ORIGINAL SOURCE:
28
29 US-08-722-626B-2

```

Query Match	25.1%;	Score 370.5;	DB 2;	Length 278;
Best Local Similarity	31.3%;	Pred. No. 9.8e-31;		
Matches 88;	Conservative 49;	Mismatches 97;	Indels 47;	Gaps 8;

```

Qy      69 FEODEVDLLRLHLKLGKRWMSLLIAGRLPGRTANDVQYNTWNTLRLKNTTKYVREKLYNN 128
      100 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      70 FTKEBETTLIOLHEMIGNRMSALIAKLTGERTDNEIKNNWHTHLKKKLDKYKPPONSIRHS 129
      100 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      129 KGEISTKEIILKPKRKAFSSSTGMGVNNNN--VLDEEHECKEIISEKOTPDASMNVND 186
      130 KSKNDHDK-----GPTTSESSNSDULTIINQKIDBPVLAAPSQTSST-- 175
Db      187 PWMINILLENCDNDEIEDEEVIVNVEKTLTSLHBEISPLNIGEGNSMOQOQISHE--NWG 245
      176 -----EMSTVTLVDDHQMVVIKQEWNES--SEYFPEID-----ESFMTDELTTDNMS 221
Qy      246 -----EFSNLTPMQOQVQNDDEPSA-----EIDLM 270
Db      222 STDHVVAANQELQVQLPSSFKRENVAIILATKQEDDDDFW 262

```

```

1  RESULT 4
2  US-08/997-251-2
3  Sequence 2, Application US/08997251
4  Patent No. 6271440
5  GENERAL INFORMATION:
6  APPLICANT: GUEBLER, FRANZ J.
7  APPLICANT: JACOBSEN, JOHN V.
8  TITLE OF INVENTION: PLANT REGULATORY PROTEINS II
9  NUMBER OF SEQUENCE: 17
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
12 STREET: 5370 Manhattan Boulevard
13 CITY: Boulder
14 STATE: CO
15 COUNTRY: US
16 ZIP: 80303
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: Patent In Release #1.0, Version #1.30
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/08/997, 251

```

```

; FILING DATE: 23-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO AU96/00383
; FILING DATE: 21-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PN6470/95
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PN3779/95
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P.
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 110-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-997-251-2

```

```

Query Match      21.5%; Score 317.5; DB 3; Length 553;
Best Local Similarity 28.2%; Pred. No. 9.7e-25;
Matches 90; Conservative 40; Mismatches 106; Indels 83; Gaps 10;

```

```

QY 11 VRKGSWTEDEDFILRKICIDYGEKWLVPFIRAGLNCRKSCRLRWLNYLRPHIKKGFDE 70
DB 42 LKKGPTMSADALIVYKKGEGNNWAVQKNTGLFRGCGSCRLRWANHLRPNLKKGAFT 101
QY 71 QDEVDLILRLHLKLNKWSLIAGRLPGRANDVKNYNTMLR-KLNTTKIVPREKINN- 128
DB 102 PEEERLILQLHSGKMGKMAAHLPGRTDNEIKYNTIRIKCORAGLPIYPASVNCOS 161
QY 129 -----KGEISTKIEIKPQRKYFSSITMGVNTNNVILDEEHCKEIISKQ- 176
DB 162 SNEDQSSDFNCGE-----NLSDDLNGNGLYLPDFTCDNFIANSEA 204
QY 177 -----TPDAS-----MDNV-----DPWVINLENCNDIDEEDEE 205
DB 205 LSYAPOLSAVISISLIGSPASKKRCGMDVNOAGMLKOSDPLPLGLSDPTINGALSSVDQ 264
QY 206 VINYEKTLTSL-----LHEEISPLNIGEGNSMOQGOISHEWNGEFS-----LNL 251
DB 265 FSNDEKTLKQALGFVDYHLEANSSSKIIAPFGALTGSHAFLN-GTFSTSRITNGPLKML 323
QY 252 PPMQOGVQNDPFAEIDLW 270
DB 324 PSLQ-----DTESDPNSW 336

```

```

RESULT 5
; US-08-997-251-4
; Sequence 4, Application US/08997251
; Patent No. 6271440
; GENERAL INFORMATION:
; APPLICANT: GUBLER, FRANZ J.
; APPLICANT: JACOBSEN, JOHN V.
; TITLE OF INVENTION: PLANT REGULATORY PROTEINS III
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Boulevard
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,251
; FILING DATE: 23-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO AU96/00383
; FILING DATE: 21-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PN6470/95
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PN3779/95
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P.
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 110-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-997-251-4

```

```

Query Match      21.3%; Score 315; DB 3; Length 553;
Best Local Similarity 35.9%; Pred. No. 1.8e-24;
Matches 65; Conservative 26; Mismatches 58; Indels 32; Gaps 3;

```

```

QY 11 VRKGSWTEDEDFILRKICIDYGEKWLVPFIRAGLNCRKSCRLRWLNYLRPHIKKGFDE 70
DB 40 LKKGPTMSADALIVYKKGEGNNWAVQKNTGLFRGCGSCRLRWANHLRPNLKKGAFT 99
QY 71 QDEVDLILRLHLKLNKWSLIAGRLPGRANDVKNYNTMLR-KLNTTKIVPREKINN- 130
DB 100 AEEERLILQLHSGKMGKMAAHLPGRTDNEIKYNTIRIK-----C 143
QY 131 GEISTKIEIKPQRKYFSSITMGVNTNNVILDEEHCKEIISKQ-----QTPASMD 183
DB 144 QRAGLPI-----YPTSVCSNEDQSSDFDCGSLNDLNLNGLYLPDFTCD 194
QY 184 N 184
DB 195 N 195

```

```

RESULT 6
; US-09-402-929-2
; Sequence 2, Application US/09402929
; Patent No. 6410825
; GENERAL INFORMATION:
; APPLICANT: Temple University - Of The Commonwealth System of Higher Education
; APPLICANT: Toscani, Antonio
; APPLICANT: Hacton, Kimi
; APPLICANT: Reddy, E. P.
; TITLE OF INVENTION: A-m/yb NULL MUTANT TRANSGENIC ANIMALS AND
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,929
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/06896
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 6056-214 PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
;
;
US-09-402-929-2

```

Query Match Similarity 16.4%; Score 242.5; DB 4; Length 75;  
Best Local Similarity 43.7%; Pred. No. 1,le-16;  
Matches 45; Conservative 19; Mismatches 38; Indels 1; Gaps 1;

Dy 13 KGSWTEDEDFLLRKCDIKXYGEGKWHLVPIRAGLNCRCKSCRLMNYLRIPIHKGDEBOD 72  
| | : | : | : | : | : | : | : | : | : | : | : | :  
Db 87 KQPMTKEDORAVETVOKYGPKRMSLI -AKHLGRIGIQCQRERWHNHLNEPVKKSWTEE 145

Oy 73 EVDLLIRKLKLGNRWSIAGRLPGRANDVNKYNNNNNLRLK 115  
| | : | | | | | : | | | | | : | : | : | : | :  
Db 146 EDRIYEARKLGNRWAEIAKLPLGRDINSIKHWNSTMRKV 188

RESULT 7  
US-09-402-929-6  
Sequence 6, Application US/09402929  
Patent No. 6410825  
GENERAL INFORMATION:  
APPLICANT: Temple University - Of The Commonwealth System of Higher Education  
APPLICANT: Toscani, Antonio  
APPLICANT: Harton, Kimi  
APPLICANT: Reddy, E. P.  
TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEIDEL, GONDA, LAVORIGNA & MONACO, P.C.  
STREET: Suite 1800 Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/402,929  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA: PCT/US98/06896  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 6056-214 PC

```

? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (215) 568-8383
? TELEFAX: (215) 568-5549
? INFORMATION FOR SEQ ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 752 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
?
?
US-03-402-929-6

```

	Query March	16.4%	Score 242.5;	DB 4;	length 752;
	Match Local Similarity	43.7%;	Pred. No.1.1e-16;		
	Batches	45;	Conservative	19;	Mismatches 38; Indels 1; Gaps 1
Oy	13 KGSWTDEDFLLRCDYKDYGEKKHVLVIRAGLNRCSRSLRWMLNYPRIHKRGDFFED	72			
	: : : :   : : : :   : : : :				
Db	87 KGPMTKEDDOVIETLVQRTGEPKRMSLT AKGLKGRLIGQCERMMNHNLNPEVKSSWTEE	145			
	: : : :   : : : :   : : : :				
Oy	73 EVDILRLHLKLGNRWSLIAGRLPGRTANDVKNVYNNTLLRKTL	115			
	: : : :   : : : :   : : : :				
Db	146 EDRIITYEAKRLGNRMALIALLLPORTDNSTIKMHNSTIMRRKV	188			

```

      RESULT 8
      US-08-928-941D-4
      ; Sequence 4, Application US/08928941D
      ; Patent No. 6180763
      ; GENERAL INFORMATION:
      ; APPLICANT: Hirai, Hiroshi
      ; APPLICANT: Sherr, Charles
      ; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES
      ; TITLE OF INVENTION: THEREOF
      ; NUMBER OF SEQUENCES: 36
      ; CORRESPONDENCE ADDRESS:
      ; STREET: 411 Hackensack Ave, Continental Plaza, 4th
      ; FLOOR: Floor
      ; CITY: Hackensack
      ; STATE: New Jersey
      ; COUNTRY: USA
      ; ZIP: 07601
      ; COMPUTER READABLE FORM:
      ; MEDIUM TYPE: Floppy disk
      ; COMPUTER: IBM PC compatible
      ; OPERATING SYSTEM: PC-DOS/MS-DOS
      ; SOFTWARE: PatentIn Release #1.0, Version #1.30
      ; CURRENT APPLICATION DATA:
      ; APPLICATION NUMBER: US/08/928,941D
      ; FILING DATE:
      ; CLASSIFICATION: 435
      ; ATTORNEY/AGENT INFORMATION:
      ; NAME: Jackson Esq., David A.
      ; REGISTRATION NUMBER: 26,742
      ; REFERENCE/DOCKET NUMBER: 1340-1-002 N CIP
      ; TELECOMMUNICATION INFORMATION:
      ; TELEPHONE: 201-487-5800
      ; TELEFAX: 201-343-1684
      ; INFORMATION FOR SEQ ID NO: 4:
      ; SEQUENCE CHARACTERISTICS:
      ; LENGTH: 156 amino acids
      ; TYPE: amino acid
      ; STRANDEDNESS:
      ; TOPOLOGY: linear
      ; MOLECULE TYPE: protein
      ; HYPOTHETICAL: NO
      ; FRAGMENT TYPE: internal
      ; ORIGINAL SOURCE:
      ; ORGANISM: Mus musculus
      ; US-08-928-941D-4
      Query Match 16.3%; Score 240.5; DB 3; Length 156
      Best Local Similarity 43.1%; Pred. No. 1.9e-17;

```

Matches 45; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

Qy 13 KGSWTDEDFLRLKCIDKYGEKMWLVPIRAGLNCRKSCRLMWLYLRPHIKRGDEPD 72  
| | | | | : | | : | | : | | : | | : | | : | | : | | : | | :  
Db 55 KGPWTKEEDORVILVOKYGPKRWSVI-AKHLKGRIGKCRERHNNHNLNPEVKTSWTEE 113  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Qy 73 EVDLILRLKHLGNRWSLIAGRLPGRTANDVKRYWNTNLLRL 115  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 114 EDRIIYQAHKRLGNRMAEIAKLPGRTDNLKNNWSTWRRKV 156  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 9  
US-09-280-590A-4  
; Sequence 4, Application US/09280590A  
; Patent No. 6303772  
; GENERAL INFORMATION:  
; APPLICANT: Hirai, Hiroshi  
; Sheri, Charles  
; Inoue, Kazushi  
; Bodner, Sarah M.  
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES  
; THEREOF  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/280,590A  
; FILING DATE: 29-Mar-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1340-1-002 N CP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 156 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; ORGANISM: Mus musculus  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-280-590A-4

Query Match 16.3%; Score 240.5; DB 4; Length 156;  
Best Local Similarity 43.7%; Pred. No. 1.9e-17;  
Matches 45; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

Qy 13 KGSWTDEDFLRLKCIDKYGEKMWLVPIRAGLNCRKSCRLMWLYLRPHIKRGDEPD 72  
| | | | | : | | : | | : | | : | | : | | : | | : | | : | | :  
Db 55 KGPWTKEEDORVILVOKYGPKRWSVI-AKHLKGRIGKCRERHNNHNLNPEVKTSWTEE 113  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Qy 73 EVDLILRLKHLGNRWSLIAGRLPGRTANDVKRYWNTNLLRL 115  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 114 EDRIIYQAHKRLGNRMAEIAKLPGRTDNLKNNWSTWRRKV 156  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 10  
US-08-928-941D-36  
; Sequence 36, Application US/08928941D  
; Patent No. 6180763  
; GENERAL INFORMATION:  
; APPLICANT: Hirai, Hiroshi  
; Sheri, Charles  
; Inoue, Kazushi  
; Bodner, Sarah M.  
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES  
; THEREOF  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,941D  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1340-1-002 N CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 156 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM: Gallus gallus  
US-08-928-941D-36

Query Match 16.0%; Score 236.5; DB 3; Length 156;  
Best Local Similarity 42.7%; Pred. No. 4.9e-17;  
Matches 44; Conservative 21; Mismatches 37; Indels 1; Gaps 1;

Qy 13 KGSWTDEDFLRLKCIDKYGEKMWLVPIRAGLNCRKSCRLMWLYLRPHIKRGDEPD 72  
| | | | | : | | : | | : | | : | | : | | : | | : | | : | | :  
Db 55 KGPWTKEEDORVILVOKYGPKRWSVI-AKHLKGRIGKCRERHNNHNLNPEVKTSWTEE 113  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Qy 73 EVDLILRLKHLGNRWSLIAGRLPGRTANDVKRYWNTNLLRL 115  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 114 EDRIIYQAHKRLGNRMAEIAKLPGRTDNLKNNWSTWRRKV 156  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 11  
US-09-280-590A-46  
; Sequence 46, Application US/09280590A  
; Patent No. 6303772  
; GENERAL INFORMATION:  
; APPLICANT: Hirai, Hiroshi  
; Sheri, Charles  
; Inoue, Kazushi  
; Bodner, Sarah M.  
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES  
; THEREOF  
; NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/280,590A  
FILING DATE: 29-Mar-1999  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1340-1-002 N CP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 156 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEetical: YES  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
ORGANISM: Gallus gallus  
SEQUENCE DESCRIPTION: SEQ ID NO: 46:  
US-09-280-590A-46  
Query Match 16.0%; Score 236.5; DB 4; Length 156;  
Best Local Similarity 42.7%; Pred. No. 4.9e-17;  
Matches 44; Conservative 21; Mismatches 37; Indels 1; Gaps 1;  
QY 13 KGSWTDEEDFLRKCDIDYGEKWHVPIRAGLNRCKSCRLRWLYLRPHIKRDEEOD 72  
DB 55 KGPWTKEEDQVIVLVQKYGPKRSVTAKLKGRIGQCKRHHNHLPVKKTSWTEE 113  
QY 73 EVDLILRLHLKLLGRMSLIAGRLPGRTANDVKNYNTNLARKL 115  
DB 114 EDRITVQAHKRLGNRMALIALLPGRFDNAIKHNSWTRRKV 156  
RESULT 12  
US-09-156-316-1  
Sequence 1, Application US/09156316  
Patent No. 6183961  
GENERAL INFORMATION:  
APPLICANT: Bernstein, Harold S.  
APPLICANT: Coughlin, Shaun R.  
TITLE OF INVENTION: Methods and Compositions for Regulating Cell Cycle  
FILE REFERENCE: UCSF-020/01US  
CURRENT APPLICATION NUMBER: US/09/156,316  
CURRENT FILING DATE: 1998-09-18  
EARLIER APPLICATION NUMBER: 60/060,688  
EARLIER FILING DATE: 1997-09-22  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 1  
LENGTH: 802  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-156-316-1

Query Match 10.7%; Score 157.5; DB 3; Length 802;  
Best Local Similarity 29.4%; Pred. No. 1e-07;  
Matches 52; Conservative 32; Mismatches 62; Indels 31; Gaps 9;  
QY 6 MSSIGYRKSGWTBDEEDFLRKCDIDYGEKWHVPIRAGLNRCKSCRLRWLYLRPH 63  
DB 1 MPRIMTKGVTWRTEDDEILAAVAKYGNQWSI--ASLLHRSKAKQCARVYEWLDP 57  
QY 64 IKRGDFEODEVDLILRLHLKLLGRMSLIAGRLPGRTANDVKNYNTNLARKL-----NT 117  
DB 58 IKRTWSREEREKTLHLAKLMPYQWRTIA-PIIGRTAQCLEHYEF-LIDKAKQRDNEE 115  
QY 118 TKIVPEKKNKCGEISTKIEIIPKPRKYSSTMMKNVNNVILDEEHKEIIE 174  
DB 116 TTDDPRKL---KGEIDPNPE-TKPAR-----PDPIDMEDE--LEMLSE 154  
RESULT 13  
US-08-519-103-16  
Sequence 16, Application US/08519103  
Patent No. 5733730  
GENERAL INFORMATION:  
APPLICANT: delange, Titia  
TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND  
METHOD OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE THEREOF  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KLAUBER & JACKSON  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/519,103  
FILING DATE: 25-AUG-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Crane-Feury, Sharon E.  
REGISTRATION NUMBER: 36,113  
REFERENCE/DOCKET NUMBER: 600-1-142  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 51 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-519-103-16  
Query Match 9.3%; Score 138; DB 1; Length 51;  
Best Local Similarity 47.1%; Pred. No. 2.1e-07;  
Matches 24; Conservative 11; Mismatches 16; Indels 0; Gaps 0;  
QY 64 IKRGDFEODEVDLILRLHLKLLGRMSLIAGRLPGRTANDVKNYNTNLARKL 114  
DB 1 VKSSWTEEDRITVYEAHKLGNRMALIALLPGRFDNSIKHNSWTRRK 51  
RESULT 14  
US-08-519-103-17  
Sequence 17, Application US/08519103  
Patent No. 5733730  
GENERAL INFORMATION:

```

; APPLICANT: delange, Titia
; TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/519,103
; FILING DATE: 25-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon B.
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 600-1-142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-519-103-17

Query Match          9.3%; Score 138; DB 1; Length 51;
Best Local Similarity 49.0%; Pred. No. 2.1e-07;
Matches 25; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY      64 IKRGDFQDEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLRLK 114
DB      1 VKKSWTEEDRITICEAHKVLGNRMWELAKWLPRTDNVAVNHNSTIKRK 51

RESULT 15
US-09-018-635-16
; Sequence 16, Application US/09018635
; Patent No. 6297356
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Broccoli, Dominique
; APPLICANT: Smogorzewska, Agata
; TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE THEREOF
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,635
; FILING DATE:
; CLASSIFICATION:
```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: David A. Jackson
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-142 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-018-635-16

Query Match          9.3%; Score 138; DB 3; Length 51;
Best Local Similarity 47.1%; Pred. No. 2.1e-07;
Matches 24; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY      64 IKRGDFQDEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLRLK 114
DB      1 VKKSWTEEDRITICEAHKVLGNRMWELAKWLPRTDNVAVNHNSTIKRK 51

Search completed: January 29, 2004, 20:13:54
Job time : 22 secs
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model ~

Run on: January 29, 2004, 20:12:52 ; Search time 33 Seconds  
(without alignments)  
1725.570 Million cell updates/sec

Title: US-10-033-190-2  
Perfect score: 1477  
Sequence: 1 MNSTMSLSLGVKRSWTDEE.....QQGVNDPFAEIDMLD 274

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues  
Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep:\*
  - 2: /cgn2\_6/prodata/1/pubpaa/PCr\_NEW\_PUB.pep:\*
  - 3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
  - 4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep:\*
  - 5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
  - 6: /cgn2\_6/prodata/1/pubpaa/PCrUS\_PUBCOMB.pep:\*
  - 7: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
  - 8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep:\*
  - 9: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep:\*
  - 10: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep:\*
  - 11: /cgn2\_6/prodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
  - 12: /cgn2\_6/prodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
  - 13: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep:\*
  - 14: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep:\*
  - 15: /cgn2\_6/prodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
  - 16: /cgn2\_6/prodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
  - 17: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
  - 18: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1477	100.0	274	US-10-033-190-2	Sequence 2, Appl1
2	542.5	36.7	246	US-09-934-455-482	Sequence 482, App
3	532.5	36.1	249	US-09-934-455-14	Sequence 14, Appl
4	532.5	36.1	249	US-10-285-403-44	Sequence 44, Appl
5	524	35.5	212	US-09-934-455-438	Sequence 438, App
6	518.5	35.1	209	US-09-934-455-480	Sequence 480, App
7	423.5	28.7	249	US-10-286-264-148	Sequence 148, Appl
8	408	27.6	269	US-09-533-029-94	Sequence 94, Appl
9	408	27.6	269	US-10-278-173-72	Sequence 72, Appl
10	401	27.1	203	US-09-934-455-390	Sequence 390, App
11	401	27.1	203	US-10-407-920-29	Sequence 29, Appl
12	397	26.9	342	US-09-934-455-252	Sequence 252, App
13	396.5	26.8	302	US-09-443-704-12	Sequence 12, Appl
14	396.5	26.8	302	US-10-008-118A-12	Sequence 12, Appl
15	394	26.7	363	US-10-093-837-2	Sequence 2, Appl1

16	393	26.6	120	9	US-09-443-704-6	Sequence 6, Appl1
17	393	26.6	120	14	US-10-008-118A-6	Sequence 6, Appl1
18	387.5	26.2	253	9	US-09-443-704-28	Sequence 28, Appl
19	387.5	26.2	253	14	US-10-008-118A-28	Sequence 28, Appl
20	385	26.1	246	9	US-09-443-704-32	Sequence 32, Appl
21	385	26.1	246	14	US-10-008-118A-32	Sequence 32, Appl
22	379.5	25.7	258	9	US-09-443-704-4	Sequence 4, Appl1
23	379.5	25.7	258	14	US-10-008-118A-4	Sequence 4, Appl1
24	377.5	25.6	274	16	US-10-278-536-36	Sequence 36, Appl
25	377	25.5	268	9	US-09-443-704-34	Sequence 34, Appl
26	377	25.5	268	14	US-10-008-118A-34	Sequence 34, Appl
27	377	25.5	349	15	US-10-278-173-138	Sequence 138, App
28	377	25.5	349	16	US-10-278-536-114	Sequence 114, App
29	374	25.3	273	9	US-09-443-704-20	Sequence 20, Appl
30	374	25.3	273	14	US-10-008-118A-20	Sequence 20, Appl
31	374	25.3	371	11	US-09-533-029-86	Sequence 86, Appl
32	374	25.3	371	11	US-09-934-455-66	Sequence 66, Appl
33	374	25.3	371	16	US-10-278-536-40	Sequence 40, Appl
34	373.5	25.3	295	15	US-10-278-173-152	Sequence 152, App
35	373.5	25.3	335	15	US-10-295-403-58	Sequence 58, Appl
36	373	25.3	281	15	US-10-295-403-46	Sequence 46, Appl
37	373	25.3	282	15	US-10-225-068-10	Sequence 10, Appl
38	371	25.1	332	15	US-10-021-811-54	Sequence 54, Appl
39	370	25.1	149	9	US-09-443-704-30	Sequence 30, Appl
40	370	25.1	149	14	US-10-008-118A-30	Sequence 30, Appl
41	370	25.1	245	15	US-10-295-403-60	Sequence 60, Appl
42	368.5	24.9	333	15	US-10-278-173-146	Sequence 146, App
43	368	24.9	120	9	US-09-443-704-50	Sequence 50, Appl
44	368	24.9	120	14	US-10-008-118A-50	Sequence 50, Appl
45	368	24.9	305	9	US-09-443-704-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1  
US-10-033-190-2  
Sequence 2, Application US/10033190  
Publication No. US20020133848A1  
GENERAL INFORMATION: Exelixis Plant Sciences, Inc.  
APPLICANT: Exelixis Plant Sciences, Inc.  
TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF AN ANTHOCYANIN MUTANT (ANTHOCYANIN)  
FILE REFERENCE: EP01-002C  
CURRENT APPLICATION NUMBER: US/10/033,190  
CURRENT FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: US 60/244,685  
PRIOR FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 274  
TYPE: PRT  
ORGANISM: Lycopersicon esculentum  
US-10-033-190-2

Query Match 100.0%; Score 1477; DB 14; Length 274;  
Best Local Similarity 100.0%; Pred. No. 2,3e-135;  
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNSTMSLSLGVKRSWTDEEFLKRCIDKYGEKWHVPIRAGLNCRCSCRRLRWNTL 60  
DB 1 MNSTMSLSLGVKRSWTDEEFLKRCIDKYGEKWHVPIRAGLNCRCSCRRLRWNTL 60  
QY 61 RPHIKRDPFQDEVDLILRLHLKLGNSWLSIAGRLPRTANDVKNYNTNLRLKNTTKI 120  
DB 61 RPHIKRDPFQDEVDLILRLHLKLGNSWLSIAGRLPRTANDVKNYNTNLRLKNTTKI 120  
QY 121 VPRKINKKGEISTKIEIKPQRKRFSSMTKNTNNNTLDEEHCCKIISKORPPA 180  
DB 121 VPRKINKKGEISTKIEIKPQRKRFSSMTKNTNNNTLDEEHCCKIISKORPPA 180  
QY 181 SMDVDPWMIINLLENCNDIEDEEVINYEKTLTSLHFEIISPPLNIGNSWQOQOIS 240

Db 181 SMDVDPWMTINLNCNDIEDEEVINYKTLTSLHEISPLNIGENSGMOQOIS 240  
Qy 241 HENWGEFSLNLPWMOGVQNDPFAEIDLMNLLD 274  
Db 241 HENWGEFSLNLPWMOGVQNDPFAEIDLMNLLD 274

## RESULT 2

US-09-934-455-482  
Sequence 482, Application US/09934455  
Publication No. US20030121070A1  
GENERAL INFORMATION:  
APPLICANT: Creelmann, Robert  
APPLICANT: Adam, Luc  
APPLICANT: Dubbell, Arnold  
APPLICANT: Heard, Jacqueline  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Keddie, James  
APPLICANT: Pilgrim, Marsha  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Reuber, Lynne  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Yu, Guo-Liang  
APPLICANT: Pineda, Omalta  
TITLE OF INVENTION: Genes for Modifying Plant Traits IV  
FILE REFERENCE: MBI-0025  
CURRENT APPLICATION NUMBER: US/09/934,455  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/227439  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: MBI-0022  
PRIOR FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: MBI-0023  
PRIOR FILING DATE: 2001-04-17  
NUMBER OF SEQ ID NOS: 516  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 482  
LENGTH: 246  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-934-455-482

Query Match 36.7%; Score 542.5; DB 11; Length 246;  
Best Local Similarity 52.4%; Pred. No. 1.7e-44;  
Matches 109; Conservative 26; Mismatches 52; Indels 21; Gaps 4;

Qy 7 SSLGVRKGSWTDDEDFLLKCIDKYGEGKWHLPTRAGLNCRRKSCRLMWNTLRLPHIKR 66  
Db 4 SPKGLRKGTWTTBEDILRCIDKYGEGKWHVPLRTGLNCRKSCRLMWNTLRLPHIKR 63  
Qy 67 GDFEODEVDLLRLHKLGNRWSLIAGRLPRTANDVKNVNTNLRLKLTN---TTKIVP 122  
Db 64 GKLCSDEVDVLRHLKLGNRWSLIAGRLPRTANDVKNVNTNLRLKLGDRCKTKYI- 122  
Qy 123 REKINNKGESTKTEIRKPKRKYFSSTMKQVT-----NNNVILDEBEH 167  
Db 123 NKNITSHTSSAQKIDVLRKPRPSDKNSCNDVILPKVDVPLHLGLNNVYCESSIT 182  
Qy 168 C-KEIISEKQTPDASMDVDPWMTINLLE 194  
Db 183 CNKDEQKDKLININLLDGDNMWESLLE 210

## RESULT 3

US-09-934-455-14  
Sequence 14, Application US/09934455  
Publication No. US20030121070A1  
GENERAL INFORMATION:  
APPLICANT: Adam, Luc  
APPLICANT: Creelmann, Robert  
APPLICANT: Dubbell, Arnold  
APPLICANT: Heard, Jacqueline

APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Keddie, James  
APPLICANT: Pilgrim, Marsha  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Reuber, Lynne  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Yu, Guo-Liang  
APPLICANT: Pineda, Omalta  
TITLE OF INVENTION: Genes for Modifying Plant Traits IV  
FILE REFERENCE: MBI-0025  
CURRENT APPLICATION NUMBER: US/09/934,455  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/227439  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: MBI-0022  
PRIOR FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: MBI-0023  
PRIOR FILING DATE: 2001-04-17  
NUMBER OF SEQ ID NOS: 516  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 14  
LENGTH: 249  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-934-455-14

Query Match 36.1%; Score 532.5; DB 11; Length 249;  
Best Local Similarity 44.4%; Pred. No. 1.6e-43;  
Matches 122; Conservative 34; Mismatches 76; Indels 43; Gaps 7;

Qy 7 SSLGVRKGSWTDDEDFLLKCIDKYGEGKWHLPTRAGLNCRRKSCRLMWNTLRLPHIKR 66  
Db 4 SSKGLRKGAWTAEDSLLRLCIDKYGEGKWHVPLRAGLNCRRKSCRLMWNTLRLPHIKR 63  
Qy 67 GDFEODEVDLLRLHKLGNRWSLIAGRLPRTANDVKNVNTNLRLKLTN---TTKIVPEKI 126  
Db 64 GRLSNDEVDLLRLHKLGNRWSLIAGRLPRTANDVKNVNTNLRLKLGDRCKTKYI- 123  
Qy 127 NNKCGEIST---KLEIRKPKRKYFSSTMKQVTNNNVILDEBEHCKEIISEKQTPDASMD 183  
Db 124 KNISPTTVOKIGVKKPRPSF---SVNNGCSHLNGLREVLLIPSCGLKKK----- 174  
Qy 184 NVDPMWNTINLEN---CNDIEDEEVINYKTLTSLHEISPLNIGENSGMOQOIS 240  
Db 175 -----NVCENSITCKDEKDFV-----NNLWNGDMWNTINL 208  
Qy 241 HENWGEFSLNLPWMOGVQNDPFAEID-LMNLLD 274  
Db 209 GEN-QEADALVPBATTAGATLAFDVEQLMSLFD 242

## RESULT 4

US-10-295-403-44  
Sequence 44, Application US/10295403  
Publication No. US20030101481A1  
GENERAL INFORMATION:  
APPLICANT: Heard, Jacqueline  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Adam, Luc  
APPLICANT: Broun, Pierre  
APPLICANT: Pineda, Omalta  
APPLICANT: Reuber, Lynne  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Keddie, James  
APPLICANT: Zhang, James  
APPLICANT: Benito, Maria-Ines  
APPLICANT: Yu, Guo-Liang  
APPLICANT: Fromm, Mike  
TITLE OF INVENTION: PLANT GENE SEQUENCES I  
FILE REFERENCE: MBI-0003  
CURRENT APPLICATION NUMBER: US/10/295,403  
CURRENT FILING DATE: 2002-11-15  
PRIOR APPLICATION NUMBER: US/09/394,519

;; PRIOR FILING DATE: 1999-09-13  
;; PRIOR APPLICATION NUMBER: 60/101,349  
;; PRIOR FILING DATE: 1998-09-22  
;; PRIOR APPLICATION NUMBER: 60/103,312  
;; PRIOR FILING DATE: 1998-10-06  
;; PRIOR APPLICATION NUMBER: 60/108,734  
;; PRIOR FILING DATE: 1998-11-17  
;; PRIOR APPLICATION NUMBER: 60/113,409  
;; PRIOR FILING DATE: 1998-12-22  
;; NUMBER OF SEQ ID NOS: 170  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO 44  
;; LENGTH: 249  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis thaliana  
;; FEATURE:  
;; OTHER INFORMATION: G663  
US-10-295-403-44

Query Match 36.1%; Score 532.5; DB 15; Length 249;  
Best Local Similarity 44.4%; Pred. No. 1.6e-43;  
Matches 122; Conservative 34; Mismatches 76; Indels 43; Gaps 7;

QY 7 SSIGVKGSTWDEEDFLRKCIDKYGEGKWHLPVIRAGLNCRCKSCRLRWNTLTPHRIKR 66  
DB 4 SSKGLRKGAWTAEEDSLRLCIDKYGEGKWHQVPLRAGLNCRCKSCRLRWNTLTPHRIKR 63  
QY 67 GDFEODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKYWNNTLRLKNTTKIVPREKI 126  
DB 64 GRLSDVBDLLRLHKLGNRWSLIAGRLPGRTANDVKYWNNTLRLKNTTKIVPREKI 123  
QY 127 NNKCGEIST--KIEIKPQRKRYFSTMKVNTNNVILDEEBHCKEIISEKQTPDASMD 183  
DB 124 KNISPTTVOKIGVKEPRSP--SVNNGSHANGLEVLDLBSCLGLKKN----- 174  
QY 184 NVDPWMINLEN--CNDIEDEEVVINYKTLTSLHBEISPPINIGESNMQOQIS 240  
DB 175 -----NVCENSITCKDKDEKDFV-----NLLMNGDNMWTLENL 208  
QY 241 HEMWGEFSLNLPMMQGVQNDPSAED-LMNLID 274  
DB 209 GEN-QEADALVPBATTAEHATLAFDVEQLMSLFD 242

RESULT 5  
US-09-934-455-438  
;; Sequence 438, Application US/09934455  
;; Publication No. US20030121070A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Adam, Luc  
;; APPLICANT: Creelman, Robert  
;; APPLICANT: Dubell, Arnold  
;; APPLICANT: Heard, Jacqueline  
;; APPLICANT: Jiang, Cai-Zhong  
;; APPLICANT: Keddle, James  
;; APPLICANT: Pilgrim, Marsha  
;; APPLICANT: Ratcliffe, Oliver  
;; APPLICANT: Reuber, Lynne  
;; APPLICANT: Riechmann, Jose Luis  
;; APPLICANT: Yu, Guo-Liang  
;; APPLICANT: Pineda, Omalra  
;; TITLE OF INVENTION: Genes for Modifying Plant Traits IV  
;; FILE REFERENCE: MBI-0025  
;; CURRENT APPLICATION NUMBER: US/09/934,455  
;; CURRENT FILING DATE: 2001-08-22  
;; PRIOR APPLICATION NUMBER: 60/227439  
;; PRIOR FILING DATE: 2000-08-22  
;; PRIOR APPLICATION NUMBER: MBI-0022  
;; PRIOR FILING DATE: 2001-11-16  
;; PRIOR APPLICATION NUMBER: MBI-0023  
;; PRIOR FILING DATE: 2001-04-17  
;; NUMBER OF SEQ ID NOS: 516  
;; SOFTWARE: Patentln version 3.1

;; SEQ ID NO 438  
;; LENGTH: 212  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis thaliana  
US-09-934-455-438

Query Match 35.5%; Score 524; DB 11; Length 212;  
Best Local Similarity 48.3%; Pred. No. 8.7e-43;  
Matches 102; Conservative 33; Mismatches 38; Indels 38; Gaps 5;

QY 7 SSIGVKGSTWDEEDFLRKCIDKYGEGKWHLPVIRAGLNCRCKSCRLRWNTLTPHRIKR 66  
DB 4 SSKGLRKGAWTAEEDSLRLCIDKYGEGKWHQVPLRAGLNCRCKSCRLRWNTLTPHRIKR 63  
QY 67 GDFEODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKYWNNTLRLKNTTKIVPREKI 114  
DB 64 GRLSDVBDLLRLHKLGNRWSLIAGRLPGRTANDVKYWNNTLRLKNTTKIVPREKI 113  
QY 115 -----LNTTKIVPREK--INNKGGEIST--KIEIKPQRKRYFSTMKVNTNN 159  
DB 124 DNPPTPALKNVYKPRPRSPFTVNNDCNHLNAPKVDVNP-----CLGLNINVCNDS 179  
QY 160 VIIDEHBCKEIISEKQTPDASMDNVDPMWT 190  
DB 180 IIVNKKKKDQLVN-----NLIDGDNMWT 203

RESULT 6  
US-09-934-455-480  
;; Sequence 480, Application US/09934455  
;; Publication No. US20030121070A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Adam, Luc  
;; APPLICANT: Creelman, Robert  
;; APPLICANT: Dubell, Arnold  
;; APPLICANT: Heard, Jacqueline  
;; APPLICANT: Jiang, Cai-Zhong  
;; APPLICANT: Keddle, James  
;; APPLICANT: Pilgrim, Marsha  
;; APPLICANT: Ratcliffe, Oliver  
;; APPLICANT: Reuber, Lynne  
;; APPLICANT: Riechmann, Jose Luis  
;; APPLICANT: Yu, Guo-Liang  
;; APPLICANT: Pineda, Omalra  
;; TITLE OF INVENTION: Genes for Modifying Plant Traits IV  
;; FILE REFERENCE: MBI-0025  
;; CURRENT APPLICATION NUMBER: US/09/934,455  
;; CURRENT FILING DATE: 2001-08-22  
;; PRIOR APPLICATION NUMBER: 60/227439  
;; PRIOR FILING DATE: 2000-08-22  
;; PRIOR APPLICATION NUMBER: MBI-0022  
;; PRIOR FILING DATE: 2001-11-16  
;; PRIOR APPLICATION NUMBER: MBI-0023  
;; PRIOR FILING DATE: 2001-04-17  
;; NUMBER OF SEQ ID NOS: 516  
;; SOFTWARE: Patentln version 3.1  
;; SEQ ID NO 480  
;; LENGTH: 209  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis thaliana  
US-09-934-455-480

Query Match 35.1%; Score 518.5; DB 11; Length 209;  
Best Local Similarity 44.0%; Pred. No. 2.9e-42;  
Matches 120; Conservative 29; Mismatches 45; Indels 79; Gaps 11;

QY 7 SSIGVKGSTWDEEDFLRKCIDKYGEGKWHLPVIRAGLNCRCKSCRLRWNTLTPHRIKR 66  
DB 4 SSKGLRKGAWTAEEDSLRLCIDKYGEGKWHQVPLRAGLNCRCKSCRLRWNTLTPHRIKR 63  
QY 67 GDFEODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKYWNNTLRLKNTTKIVPREKI 126  
DB 64 GRLSDVBDLLRLHKLGNRWSLIAGRLPGRTANDVKYWNNTLRLKNTTKIVPREKI 123

[illegible]

RESULT 7  
US-10-28

```

Sequence 148, Application US/10286264
Publication No. US2003009387A1
GENERAL INFORMATION:
APPLICANT: Keddie, James
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Ratcliffe, Oliver
APPLICANT: Zhang, James
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Pineda, Omaira
APPLICANT: Heard, Jacqueline
APPLICANT: Yu, Guo-Liang
APPLICANT: Adam, Luc
APPLICANT: Broun, Pierre
APPLICANT: Reuber, Lynne
APPLICANT: Pilgrim, Marsha
APPLICANT: Samaha, Raymond
TITLE OF INVENTION: POLYNUCLEOTIDES FOR SEED TRAIT ALTERATION
FILE REFERENCE: MB1-008
CURRENT APPLICATION NUMBER: US/10/266,264
CURRENT FILING DATE: 2002-11-01
PRIORITY APPLICATION NUMBER: 60/125,814
PRIORITY FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 165
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 148
LENGTH: 249
TYPE: PRN
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G211
US-10-266-264-148

```

	Query Match	28.7%	Score 423.5;	DB 15;	Length 249;
	Best Local Similarity	36.5%;	Pred. No. 6.5e-33;		
	Matches	100;	Conservative	32;	Mismatches 65; Indels 77; Gaps 6
Qy	7	SLGVKSGSWTDEEDFLRKICIDKYGECKMVLPIRAGINCRKSCRLMLNLYLAPHIKR	66		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Dd	19	TKMGKKRGWTVYEEDELVSFIKKGEGWRSLPRAGILRCRKSCLRMNNYLPSVXR	78		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Qy	67	GDFEODEVDLIIRLHLKLGNRWSLIAGRIPLGTANDVKYNTNTLIRLKINTKIYPREKI	126		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Dd	79	GGITSDEEDLLIRLRLLGNRWSLIAGRTPGRTDNEIKRYNTHLRKTLROGIPO---	135		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Qy	127	NNKCEISITKEIIIPQRKYSSIMKANVTNNNVILDEE-----HCKELISEKPTDA	180		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Dd	136	-----THKPLDANNIHKRPEEVSYGOKYPLEPISSSHDTDT	171		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Qy	181	SMDNVDPMMWINLE-----NCNDJIEEDDEVINYNETLSLTHEEISPLNI	228		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Dd	172	TVMGGDGSKSINSINVFGHEGYEDPGFCYD-----KSSFSLNSLIN-DVGDPFG-	220		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Qy	229	GEGNSMOOGQLSHENWGFEFSLNIPMQOQGVDND	262		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Dd	221	-----NIIPSLQPLQND	233		

## RESULT 8

US-09-533-029-94  
 ? Sequence 929, Application US/09533029  
 ? Publication No. US20030046723A1  
 ? GENERAL INFORMATION:  
 ? APPLICANT: Heard, Jacqueline  
 ? APPLICANT: Bronn, Pierre  
 ? APPLICANT: Riechmann, Jose-Luis  
 ? APPLICANT: Keddie, James  
 ? APPLICANT: Pineda, Omarla  
 ? APPLICANT: Adam, Luc  
 ? APPLICANT: Samaha, Raymond  
 ? APPLICANT: Zhang, James  
 ? APPLICANT: Yu, Guo-Liang  
 ? APPLICANT: Ratcliffe, Oliver  
 ? APPLICANT: Pilgrim, Marsha  
 ? APPLICANT: Jiang, Cai-Zhong  
 ? APPLICANT: Reuber, Lynne  
 ? TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES  
 ? FILE REFERENCE: MBI-010  
 ? CURRENT APPLICATION NUMBER: US/09/533,029  
 ? CURRENT FILING DATE: 2000-03-22  
 ? EARLIER APPLICATION NUMBER: 60/125,814  
 ? EARLIER FILING DATE: 1999-03-23  
 ? NUMBER OF SEQ ID NOS: 121  
 ? SOFTWARE: PatentIn Ver. 2.1  
 ? SEQ ID NO 94  
 ? LENGTH: 269  
 ? TYPE: PRT  
 ? ORGANISM: Arabidopsis thaliana  
 ? FEATURES:  
 ? OTHER INFORMATION: G255  
 ? US-09-533-029-94

Query Match	27.6%	Score 408	DB 11	Length 269
Best Local Similarity	39.7%	Pred. No. 2	3e-31	
Matches	96	Conservative	37	Mismatches 89
			Indels	26
			Gaps	9
QY	11	VRKSSWIDEEFPLRKCDIKDYGECKMHLVPIRAGLNRRKSCRLPFWLYRPHIRGDE	70	
Db	12	MNKAMTVEEDERLVSYLKSHEGQWRSLPPRAGLLRCKSCRLMWINLRDLRGNGTT	71	
QY	71	ODEVDLILRLKLLGNRWSLIAGRLPGRTANDVKNVWNTNLRLKNTTKIYV-REKINN	128	
Db	72	HDEDELIILKLHSLGNKMSLLAARLPGRTDNEIKVWYTHIRKLLSKGIDPATRGINE	131	
QY	129	-KCGEL-STKLEIK-----PQRKYFSTMGVNNNVIIDEEHCKEIISEKQPE	178	
Db	132	AKISDLDKTKQOIVADVSFTKFEBTDKSGQOKYIRNGLVCEE--RVEVEKIGP	188	
QY	179	DASND-NVDPWIMLLR--NC-----NDDEEDDEVIVINEKTLTSLHSEIS-PLPN	227	
Db	189	DLNLELRISPWQNRQREISTCTASRFYVENDEMCSEFVKKQTESSSISSYSSIDSSN	248	
QY	228	IG 229		
Db	249	VG 250		

## RESULT 5

```

US-10-278-173-72
; Sequence 72. Application US/10278172
; Publication No. US20030061637A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Brown, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Omatia
; APPLICANT: Pineda, Omar
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Keddie, James
; APPLICANT: Heard, Jacqueline

```

APPLICANT: Reuber, Lynne  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Adam, Luc  
APPLICANT: Samaha, Raymond  
TITLE OF INVENTION: POLYNUCLEOTIDES FOR ROOT TRAIT ALTERATION  
FILE REFERENCE: MBI-009  
CURRENT APPLICATION NUMBER: US/10/278,173  
CURRENT FILING DATE: 2002-10-21  
PRIOR APPLICATION NUMBER: US/09/533,392  
PRIOR FILING DATE: 2000-03-22  
PRIOR APPLICATION NUMBER: 60/125,814  
PRIOR FILING DATE: 1999-03-23  
NUMBER OF SEQ ID NOS: 177  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 72  
LENGTH: 269  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
OTHER INFORMATION: G255  
US-10-278-173-72

Query Match 27.6%; Score 408; DB 15; Length 269;  
Best Local Similarity 39.7%; Pred. No. 2.3e-31;  
Matches 96; Conservative 37; Mismatches 83; Indels 26; Gaps 9;

QY 11 VRKSWTDEEDFLRKCIDKYGEGKMHLPVRAGLANCRKSCRLRMNLYLRPHIKRGDFE 70  
DB 12 MNGATKEDEDLVSTIKSHGECMRSLPRAAGLRCGKSCRLRMNLYLRPHIKRGDFE 71  
QY 71 QDEVDLIRLKLHKLGNRWSLIAGRLPGRTANDVQYNNVNLRLKLTAKTIVP--REKINN 128  
DB 72 HDEDLIRLKLHKLGNRWSLIAGRLPGRTANDVQYNNVNLRLKLTAKTIVP--REKINN 111  
QY 129 -KCGEI-STKIEIITK-----PQRKRYSSMTKQNTNNVILDEEHCKEIISEKQTP 178  
DB 132 AKSIDLKTQDQIVKQVSFYTKPEETDKSGDOKNKYIRNGLVCKEE---RVVVEEKIGP 188  
QY 179 DASMD-NDVDEVMNMLE--NC-----NDIEDEEVNINYEKTLTSLHEEIS--PPLN 227  
DB 189 DLNLELRISPPWQOREISTCTJASRFYEMNDMECSSETVACQENSSSISYSSIDISSN 248  
QY 228 IG 229  
DB 249 VG 250  
RESULT 10  
US-09-934-455-390  
Sequence 390, Application US/09934455  
Publication No. US20030121070A1  
GENERAL INFORMATION:  
APPLICANT: Adam, Luc  
APPLICANT: Creelman, Robert  
APPLICANT: Dubell, Arnold  
APPLICANT: Heard, Jacqueline  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Keddie, James  
APPLICANT: Pilgrim, Marsha  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Reuber, Lynne  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Yu, Guo-Liang  
APPLICANT: Pineda, Omatia  
TITLE OF INVENTION: Genes for Modifying Plant Traits IV  
FILE REFERENCE: MBI-0025  
CURRENT APPLICATION NUMBER: US/09/934,455  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/227439  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: MBI-0022  
PRIOR FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: MBI-0023

PRIOR FILING DATE: 2001-04-17  
NUMBER OF SEQ ID NOS: 516  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 390  
LENGTH: 203  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-934-455-390

Query Match 27.1%; Score 401; DB 11; Length 203;  
Best Local Similarity 48.7%; Pred. No. 7.6e-31;  
Matches 77; Conservative 26; Mismatches 49; Indels 6; Gaps 2;

QY 12 RKGSWTDEEDFLRKCIDKYGEGKMHLPVRAGLANCRKSCRLRMNLYLRPHIKRGDFE 71  
DB 17 KKGWLTVEEDKILMDYVKAHKGHMRIAKGTGLKRCGKSCRLRMNLYLRPHIKRGDFE 76  
QY 72 DEVDLIRLKLHKLGNRWSLIAGRLPGRTANDVQYNNVNLRLKLTAKTIVPREKINNKG 131  
DB 77 QDEDLIRLKLHKLGNRWSLIAGRLPGRTANDVQYNNVNLRLKLTAKTIVPREKINNKG 132  
QY 132 EISTKIEIITKPKRRKXFFSSMTKQNTNNVILDE--EEH 167  
DB 133 DIVYQINLPNPTSETSETKISNIVDNNNIIIGDEIQEDH 170

RESULT 11  
US-10-407-920-29  
Sequence 29, Application US/10407920  
Publication No. US20040006797A1  
GENERAL INFORMATION:  
APPLICANT: Shi, Lifang  
TITLE OF INVENTION: MYB Transcription Factors and Uses for Crop Improvement  
FILE REFERENCE: 38-21(52703)A  
CURRENT APPLICATION NUMBER: US/10/407,920  
CURRENT FILING DATE: 2003-06-06  
PRIOR APPLICATION NUMBER: US/60/370,759  
PRIOR FILING DATE: 2002-04-10  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 29  
LENGTH: 203  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-10-407-920-29

Query Match 27.1%; Score 401; DB 12; Length 203;  
Best Local Similarity 48.7%; Pred. No. 7.6e-31;  
Matches 77; Conservative 26; Mismatches 49; Indels 6; Gaps 2;

QY 12 RKGSWTDEEDFLRKCIDKYGEGKMHLPVRAGLANCRKSCRLRMNLYLRPHIKRGDFE 71  
DB 17 KKGWLTVEEDKILMDYVKAHKGHMRIAKGTGLKRCGKSCRLRMNLYLRPHIKRGDFE 76  
QY 72 DEVDLIRLKLHKLGNRWSLIAGRLPGRTANDVQYNNVNLRLKLTAKTIVPREKINNKG 131  
DB 77 QDEDLIRLKLHKLGNRWSLIAGRLPGRTANDVQYNNVNLRLKLTAKTIVPREKINNKG 132  
QY 132 EISTKIEIITKPKRRKXFFSSMTKQNTNNVILDE--EEH 167  
DB 133 DIVYQINLPNPTSETSETKISNIVDNNNIIIGDEIQEDH 170

RESULT 12  
US-09-934-455-252  
Sequence 252, Application US/09934455  
Publication No. US20030121070A1  
GENERAL INFORMATION:  
APPLICANT: Adam, Luc  
APPLICANT: Creelman, Robert  
APPLICANT: Dubell, Arnold  
APPLICANT: Heard, Jacqueline  
APPLICANT: Jiang, Cai-Zhong

```

APPLICANT: Keddie, James
APPLICANT: Pilgrim, Marsha
APPLICANT: Ratcliffe, Oliver
APPLICANT: Reuber, Lynne
APPLICANT: Riechmann, Jose Luis
APPLICANT: Yu, Guo-Liang
APPLICANT: Pineda, Omalta
FILE OF INVENTION: Genes for Modifying Plant Traits IV
FILE REFERENCE: MBI-0025
CURRENT APPLICATION NUMBER: US/09/934,455
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/227439
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: MBI-0022
PRIOR FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: MBI-0023
PRIOR FILING DATE: 2001-04-17
NUMBER OF SEQ ID NOS: 516
SOFTWARE: PatentIn version 3.1
SEQ ID NO 252
LENGTH: 342
TYPE: PR1
ORGANISM: Arabidopsis thaliana
US-09-934-455-252

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 20:09:31 ; Search time 20 Seconds  
(without alignments)  
1317.511 Million cell updates/sec

Title: US-10-033-190-2

Perfect score: 1477

Sequence: 1 MNSTMSLSLGRKGSWTDEE.....QQGVQNDPFAEIDLWLLD 274

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

\* Listing first 45 summaries

Database :

PIR\_76: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	542.5	36.7	246	2 A96689	hypothetical prote
2	535	36.2	248	2 B96608	probable transcrip
3	532.5	36.1	249	2 T51687	myb-related transc
4	524	35.5	212	2 T51680	myb-related transc
5	503.5	34.1	139	2 B96689	hypothetical prote
6	423.5	28.7	249	1 S68688	myb-related protei
7	408	27.6	269	1 S58282	probable MYB fami
8	404	27.4	264	2 T03974	anthocyanin biosyn
9	403	27.3	246	2 T51631	probable transcrip
10	403	27.3	257	2 T00780	myb-related protei
11	402.5	27.3	226	2 T09773	myb-related protei
12	401	27.1	203	2 T51420	myb transcrip
13	400	27.1	193	2 T09744	myb-related protei
14	397	26.9	266	2 T03972	anthocyanin biosyn
15	397	26.9	267	2 T01188	anthocyanin biosyn
16	397	26.9	271	2 T03715	anthocyanin biosyn
17	396	26.8	198	2 T09743	myb-related protei
18	394	26.7	273	1 TVZMMB	anthocyanin biosyn
19	390.5	26.4	272	1 S69189	myb-related protei
20	389	26.3	263	2 T07395	myb-related transc
21	388	26.3	256	2 T01189	anthocyanin biosyn
22	388	26.3	325	2 T51509	probable transcrip
23	387	26.2	321	2 T02987	myb-related protei
24	386.5	26.2	274	1 T00957	myb-related protei
25	386	26.1	280	1 S26604	myb-related protei
26	379.5	25.7	232	1 U00960	myb-related protei
27	378	25.6	267	1 S61506	myb-related protei
28	378	25.6	294	2 T09879	myb-related protei
29	377.5	25.6	236	2 D85096	probable DNA-bindi

30	377.5	25.6	274	2 T05769	myb-related protei
31	377	25.5	745	2 A71448	probable MYB trans
32	375	25.4	280	2 T00737	myb-related protei
33	375	25.4	280	2 T51667	myb-related transc
34	374.5	25.4	371	2 T00438	probable MYB fami
35	374	25.3	275	2 T02988	myb-related protei
36	373.5	25.3	333	2 T45720	probable transcrip
37	373	25.3	239	2 T02984	myb-related protei
38	373	25.3	282	2 T05690	myb-related transc
39	372.5	25.2	343	2 T52590	probable transcrip
40	372.5	25.2	371	2 T51636	myb-related transc
41	372	25.2	276	2 T02985	myb-related protei
42	372	25.2	282	2 T51632	myb-related transc
43	370.5	25.1	246	1 S71283	myb-related protei
44	370.5	25.1	278	2 T03850	myb-related protei
45	370.5	25.1	333	2 T51650	probable transcrip

#### ALIGNMENTS

RESULT 1  
A96689  
hypothetical protein T27F4.12 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cross)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: A96689  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Greasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hultzer, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malti, R.; Marzali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon, I.; ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: A96689  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-246 <STO>  
A/Cross-references: GB:A005173; NID:G10092494; PIDN:AA612894.1; GSPDB:GN00141  
C/Genetics:  
A/Gene: T27F4.12  
A/Map position: 1  
Query Match 36.7%; Score 542.5; DB 2; Length 246;  
Best Local Similarity 52.4%; Pred. No. 1.3e-32;  
Matches 109; Conservative 26; Mismatches 52; Indels 21; Gaps 4;  
Qy 7 SSLGVRKGSWTDEBDFLLKRCIDKYGEGKWLHPIRAGLNRCKRSCLRLNTLRPHIKR 66  
Db 4 SPGLRKGTWTEEDILLRCIDKYGEGKWLHPIRAGLNRCKRSCLRLNTLRPHIKR 63  
Qy 67 GDEPQDEVDLILRLKILGRWSLIGRLGRANDVKNVNNILRKLN-----TTKIVP 122  
Db 64 GKLCSDVDLVLRLKILGRWSLIGRLGRANDVKNVNNILRKLN-----TTKIVP 122  
Qy 123 REKINNKGEISTKIELIKPQRKRYFSSTMKNTV-----NNNVILDEEH 167  
Db 123 NKNITSHPTSSAKIDVLRKPRPSFSDKNSCNDVNLIPKVDVPLHLGLNNVYCCSSIT 182  
Qy 168 C-KRISSEKOTPPASMDNDVPMWNLIE 194  
Db 183 CNKDEQDKLININLIDGDMNWESLIE 210  
RESULT 2  
B96608  
probable transcription factor P25P12.92 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cross)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C/Accession: B96608  
 R/Theologis: A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizart, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/Authors: Salberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A/Reference number: A86141; MUID:21016719; PMID:11130712  
 A/Accession: B96608  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-248 <STO>  
 A/Cross-references: GB:A8005173; NID:g9954749; PIDN:AAG09100.1; GSPDB:GN00141  
 C/Genetics:  
 A:Gene: F25P12.92  
 A:Map position: 1  
 C:Superfamily: Arabidopsis 28K leaf-specific myb-related protein; myb DNA-binding repeat

Query Match 36.2%; Score 535; DB 2; Length 248;  
 Best Local Similarity 44.2%; Pred. No. 4.6e-32;  
 Matches 122; Conservative 40; Mismatches 68; Indels 46; Gaps 10;  
 Db 7 SSGVRRKGSWTDEDFLLRKCIDKYGKWHVPIRAGLNCRKSCRLRWLNTLRPHIKR 66  
 4 SSKGIRKGAWTTEBDSLLRQICNKYGKWHVPIRAGLNCRKSCRLRWLNTLRPHIKR 63  
 Oy 67 GDFEODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNWNTNLTLLKNTTKIVPREKI 126  
 Db 64 GKLSSDEVLDLRLHKLGNRWSLIAGRLPGRTANDVKNWNTNLTLLKNTTKIVPREKI 118  
 Oy 127 NNKCGEIST---KLEIKPQRRKYSSTGMKNTNNVILDEBEHCKEIISEKOTPD 179  
 Db 119 KKKKADIPITPTPLKNNVYKPRPSF-----TVNN-----DCNHLNAPPKV-- 161  
 Oy 180 ASMNDVDMWNTLNLENCDDIDEBEVINYEKTLTSLHBEISPLNIGGNSMQQOIS 239  
 Db 162 ----DVNPCLGL--NINVC--DMSIITNKKDKKQQLVN-----NLIDGDMWLEKF 206  
 Oy 240 SHENWGEPSLNPWQGVQNDPSAIED-LMNLLD 274  
 Db 207 LEES-QEVDIVPEATTAHEGATLAFDVQLMSLFD 241

RESULT 3  
 T51687  
 myb-related transcription factor MYB90 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 02-Sep-2000  
 C/Accession: T51687  
 R/Kranz, H.D.; Denekamp, M.; Greco, R.; Jin, H.; Kranz, H.D.; Denekamp, M.; Greco, R.; J. Paz-Ares, J.; Weishaar, B.  
 Plant J. 16, 263-276, 1998  
 A/Title: Towards functional characterisation of the members of the R2R3-MYB gene from Ar  
 A/Reference number: Z14349; MUID:9839469; PMID:9839469  
 A/Accession: T51687  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-249 <KRA>  
 A/Cross-references: EMBL:AF062915; PIDN:AAC83637.1  
 A/Experimental source: cultivar Columbia  
 C/Genetics:  
 A:Gene: MYB90  
 C:Superfamily: Arabidopsis 28K leaf-specific myb-related protein; myb DNA-binding repeat  
 C/Keywords: transcription factor

Query Match 36.1%; Score 532.5; DB 2; Length 249;  
 Best Local Similarity 44.4%; Pred. No. 7.1e-32;  
 Matches 122; Conservative 34; Mismatches 76; Indels 43; Gaps 7;

Oy 7 SSGVRRKGSWTDEDFLLRKCIDKYGKWHVPIRAGLNCRKSCRLRWLNTLRPHIKR 66  
 Db 4 SSKGIRKGAWTTEBDSLLRQICNKYGKWHVPIRAGLNCRKSCRLRWLNTLRPHIKR 63  
 Oy 67 GDFEODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNWNTNLTLLKNTTKIVPREKI 126  
 Db 64 GKLSSDEVLDLRLHKLGNRWSLIAGRLPGRTANDVKNWNTNLTLLKNTTKIVPREKI 123  
 Oy 127 NNKCGEIST---KLEIKPQRRKYSSTGMKNTNNVILDEBEHCKEIISEKOTPDASMD 183  
 Db 124 KNIISPTTVQKIGVKKPRPSF---SVNNGSHNLGLEVDLIPSCGLAKN----- 174  
 Oy 184 NVDPWMNTLEN---CNDIDEBEVINYEKTLTSLHBEISPLNIGGNSMQQOIS 240  
 Db 175 -----NVCENSTITCKDEKDFV-----NLIANGDMWLENTL 208  
 Oy 241 HENWGEPSLNPWQGVQNDPSAIED-LMNLLD 274  
 Db 209 GEN-QEADIVPEATTAHEGATLAFDVQLMSLFD 242

RESULT 4  
 T51680  
 myb-related transcription factor MYB75 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
 C/Accession: T51680  
 R/Kranz, H.D.; Denekamp, M.; Greco, R.; Jin, H.; Kranz, H.D.; Denekamp, M.; Greco, R.; J. Paz-Ares, J.; Weishaar, B.  
 Plant J. 16, 263-276, 1998  
 A/Title: Towards functional characterisation of the members of the R2R3-MYB gene from Ar  
 A/Reference number: Z14349; MUID:9839469; PMID:9839469  
 A/Accession: T51680  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-212 <KRA>  
 A/Cross-references: EMBL:AF062908; PIDN:AAC83630.1  
 A/Experimental source: cultivar Columbia  
 C/Genetics:  
 A:Gene: MYB75  
 C:Keywords: transcription factor

Query Match 35.5%; Score 524; DB 2; Length 212;  
 Best Local Similarity 48.3%; Pred. No. 2.4e-31;  
 Matches 102; Conservative 33; Mismatches 38; Indels 38; Gaps 5;  
 Db 7 SSGVRRKGSWTDEDFLLRKCIDKYGKWHVPIRAGLNCRKSCRLRWLNTLRPHIKR 66  
 4 SSKGIRKGAWTTEBDSLLRQICNKYGKWHVPIRAGLNCRKSCRLRWLNTLRPHIKR 63  
 Oy 67 GDFEODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNWNTNLTLLKNTTKIVPREKI 114  
 Db 64 GKLSSDEVLDLRLHKLGNRWSLIAGRLPGRTANDVKNWNTNLTLLKNTTKIVPREKI 123  
 Oy 115 -----LNTTKIVPREK---INNKCGEIST---KLEIKPQRRKYSSTGMKNTNN 159  
 Db 124 DITPTPTPLKNNVYKPRPSFTVNDCHNLNAPKVDVPP-----CLGLNINNVCDNS 179  
 Oy 160 VILDEBEHCKEIISEKOTPDASMDNVDPWM 190  
 Db 180 IIVNKDKKQQLVN-----NLIDGDMWLENTL 203

RESULT 5  
 B96689  
 hypothetical protein T27F4.13 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Aug-2002  
 C/Accession: B96689  
 R/Theologis: A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizart, L.  
 Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luroso, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B96689

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-139 <STO>

A;Cross-references: GB:AE005173; NID:g10092496; PIN:AA612896.1; GSPDB:GN00141

C;Genetics:

A;Gene: T27F4.13

A;Map position: 1

C;Superfamily: Arabidopsis myb-related protein Y19; myb DNA-binding repeat homology

Query Match 34.1%; Score 503.5; DB 2; Length 139;

Best Local Similarity 70.1%; Pred. No. 4.6e-30;

Matches 96; Conservative 10; Mismatches 18; Indels 13; Gaps 2;

QY 7 SSLGVRGSGWTDEEDFLRKICIDYKGGKMHLPVIRAGLNRCRSCRLRLMNTLRPHIKR 66

DB 4 SSKGLRGAWTAEEDSLRQCTGKYGKMHQVPLRAGLNRCRSCRLRLMNTLRPHIKR 63

QY 67 GDFEODVDLILRLHKLIGNRWSLIAGRLPGRTANDVKNYNTVTLRLKNTTKIIVPEKI 126

DB 64 GKSSDVEDLILRLHKLIGNRWSLIAGRLPGRTANDVKNYNTVTLRLKNTTKIIVPEKI 111

QY 127 NNKGEISTKLEIKP 142

DB 112 HEPCCTKIKRINITP 128

RESULT 6

56868

myb-related protein 5 - *Arabidopsis thaliana*

N;Alternate names: protein Atmb5

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 15-Nov-1996 #sequence\_revision 27-Feb-1997 #text\_change 18-Jun-1999

C;Accession: S68688

R;Li, S.F.; Santini, J.M.; Nicolaou, O.; Parish, R.W.

FEBS Lett. 379, 117-121, 1996

A;Title: A novel myb-related gene from *Arabidopsis thaliana*.

A;Reference number: S68688; MUID:96225952; PMID:8635574

A;Accession: S68688

A;Molecule type: DNA

A;Residues: 1-249 <LIS>

A;Cross-references: EMBL:U26935; NID:g1254994; PIDN:AA649311.1; PID:g1218000

C;Experimental source: strain Landberg erecta

C;Genetics:

A;Gene: myb5

A;Intons: 99/3

C;Superfamily: Arabidopsis myb-related protein 5; myb DNA-binding repeat homology

C;Keywords: DNA binding; duplication; nucleus; transcription regulation

F;20-72/Domain: myb DNA-binding repeat homology <MYB1>

F;73-123/Domain: myb DNA-binding repeat homology <MYB2>

Query Match 28.7%; Score 423.5; DB 1; Length 249;

Best Local Similarity 36.5%; Pred. No. 6.8e-24;

Matches 100; Conservative 32; Mismatches 65; Indels 77; Gaps 6;

QY 7 SSLGVRGSGWTDEEDFLRKICIDYKGGKMHLPVIRAGLNRCRSCRLRLMNTLRPHIKR 66

DB 19 TKMGAKGPMVTEDEILVSEIKKEGGRWSLPKRAGLRCGSCRLRLMNTLRPHIKR 78

QY 67 GDFEODVDLILRLHKLIGNRWSLIAGRLPGRTANDVKNYNTVTLRLKNTTKIIVPEKI 126

DB 79 GGTISDEEDLILRLHKLIGNRWSLIAGRLPGRTANDVKNYNTVTLRLKNTTKIIVPEKI 135

QY 127 NNKGEISTKLEIKPORKYSSMTKNTNNVILDEE-----HCKEIISEKQTPDA 180

DB 136 -----TKRPLDANNIHKPEEVSGGQKPLPEIISSTHTDT 171

QY 181 SMDVDPWMINLE-----NCNDIEDEEVINYEKTYLSLHEEISPLNT 228

DB 172 TVNGGDSDSKNSLNVGGEHYEDPFCYVD-----KFSPLNSLIN-DVGDFFG- 220

QY 229 GEGNSMOGQGISHEMNGEFSLNPPMOGQYQND 262

DB 221 -----NIPIISQPLQMD 233

RESULT 7

558292

probable MYB family transcription factor At2g16720 [imported] - *Arabidopsis thaliana*

N;Alternate names: myb-related protein Y49

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 02-Mar-2001

C;Accession: S58292; D84543

R;Quaedvlieg, N.E.M.; Dockx, J.; Keultjes, G.G.M.; Smeekens, J.C.M.

submitted to the EMBL Data Library, July 1995

A;Reference number: S58280

A;Accession: S58292

A;Molecule type: DNA

A;Residues: 1-269 <QUA>

A;Cross-references: EMBL:X90385; NID:g928968; PIDN:CAA62033.1; PID:g928969

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentho, M.I.; Town, C.D.; Fujii, C.Y.; I M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: D84543

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-269 <STO>

A;Cross-references: GB:AE002093; NID:g4581115; PIDN:AA624605.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g16720

A;Map position: 2

A;Intons: 88/2

C;Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology

C;Keywords: DNA binding; duplication; nucleus; transcription regulation

F;9-61/Domain: myb DNA-binding repeat homology <MYB1>

F;62-112/Domain: myb DNA-binding repeat homology <MYB2>

Query Match 27.6%; Score 408; DB 1; Length 269;

Best Local Similarity 39.7%; Pred. No. 1e-22;

Matches 96; Conservative 37; Mismatches 83; Indels 26; Gaps 9;

QY 11 VRKGSWTDEEDFLRKICIDYKGGKMHLPVIRAGLNRCRSCRLRLMNTLRPHIKR 70

DB 12 MKGAMTKDEEDLVSYIKSHGSCWRSLPRAAGLRCGSCRLRLMNTLRPHIKR 71

QY 71 QDEVDLILRLHKLIGNRWSLIAGRLPGRTANDVKNYNTVTLRLKNTTKIIVPEKI 128

DB 72 HDEDELILKLSLIGNRWSLIAGRLPGRTANDVKNYNTVTLRLKNTTKIIVPEKI 131

QY 129 -KGEI-STKLEIK-----PQRKYSSMTKNTNNVILDEEHCKEIISEKQTP 178

DB 132 AKSDELKKTQIVKDVSVTKKEETDKSGDQKRYRNGKCKE-----RVVKEKIGP 188

QY 179 DASMD-NVDPMWNLLE--NC-----NDIEDEEVINYEKTYLSLHEEIS-PLN 227

DB 189 DNLIELRISFPWONOREISTCTASRFYEMDMECSSSETVACQRENSISYSSIDISSN 248

QY 228 IG 229

DB 249 VG 250

RESULT 8

T03974

anthocyanin biosynthesis regulatory protein - maize

C;Species: Zea mays (maize)

	Query Match	Score	404;	DB 2;	Length	264;
	Best Local Similarity	64.3%;	Pred.	No.1,9e-22;		
Matches	72; Conservative	17;	Mismatches	23;	Indels	0; Gaps
Oy	10	GVRGKSWTDEDFLLRKIDDKGECKHLVPRBAGLNCRRSKRLRWLYRPHIKREFD	69			
Dd	11	GKRKAMPAKEEDITLAAYVKAHBGSKREVPQAKGLRCRGCSCRMLNTNYLRNFKRSNI	70			
Oy	70	EODEVDLLIRLHKLGNGRWSLIAGLPGRYANDAVNNYNILNRKLNTTKIV	121			
Dd	71	SYDEDLLIVRLHKLGNRWSLIAGRLPRTINEINRYNNYSTLTGRAGASRVV	122			

```

RESULT 9
TS1631
probable transcription factor MYB3 [imported] - Arabidopsis thaliana
C|Species: Arabidopsis thaliana (mouse-ear cress)
C|Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000
C|Accession: TS1631
R|Kranz, H.D.; Denekamp, M.; Greco, R.; Jin, H.; Kranz, H.D.; Denekamp, M.; Greco, R.;
; Paz-Ares, J.; Weishaar, B.
Plant J. 16, 263-276, 1998
A|Title: Towards functional characterization of the members of the R2R3-MYB gene from A
A|Reference number: Z14349; MUID:9839469; PMID:9839469
A|Accession: TS1631
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: mRNA
A|Residues: 1,246 <KRA>
A|Cross-references: EMBL:AF062859; PIDD:AA083581.1
A|Experimental source: cultivar Columbia
C|Genetics:
A|Gene: MYB3
C|Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology

Query Match      27.3%; Score 403; DB 2; Length 246;
Best Local Similarity 37.3%; Pred. No. 2,1e-22;
Matches 91; Conservative 41; Mismatches 68; Indels 44; Gaps 7;

```

```

QY 11 FRKSGWDBEDBLLEKCLDKYGECKHVLVPIRAGINROCKSRMLWYLRPHIRKGEFE 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKKGAMTKEEDLDVADYIRKRGEGCWRSLPPRAGJQRCCGKSCRLLMWMYLRFDLRRGNFT 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 71 ODEVLLIRLHLGLLGRWMSLIAGRLPGRIANDVQVWYVWNTNLRKINTTKYVPR- EKINN 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 EBEDELITKHLSLGKNSKSLIGRLPGRDNEITKYYWMTMIRKLLSKGIDNSHRLINE 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 CGEISTKLEIIPQRKYYFSSITMKQVNTNNVILDEEHCK- EIISKOTPDASMDNDVPW 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 -----SVSP-----SSLQNDVETIHLDFSGVPKPRPVRBE----- 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 169 WTNLLENCND-----DIEBDEEVYNIEKTLTSLHHEISPPRANIGSGNSMOGQISH 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 -IGWANNCCSGTSEKDYGNEDWVNLNIELSVGSPSYRESTRKYSVDNSAE-----ST 205
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 10  
T00780  
myb-related protein T22J18.19 - Arabidopsis thaliana  
C1Species: Arabidopsis thaliana (mouse-ear cress)  
C1Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 22-Oct-1999  
C1Accession: T00780  
R1Vycostsaia, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Oji, O.; Kwan, A.; Liu, S.; Li,  
rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N  
submitted to the EMBL Data Library, July 1998  
A1Description: Arabidopsis thaliana chromosome 1 BAC T22J18 sequence, complete sequence.

A:Accession: T00780  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-257 <VYS>  
A:Cross-references: EMBL:AC003979; NID:g3172156; PIDN:AC0552.1; PID:g32694; GSPDB:GN000  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Gene: ATSP:T22J18.19  
A:Map position: 1  
A:Introns: 45/1; 86/2  
C:Superfamily: barley myb-DNA-related protein 1; myb DNA-binding repeat homology  
C:Keywords: DNA binding; duplication; nucleus; transcription regulation  
F:9-6/Domain: myb DNA-binding repeat homology <MYB1>  
F:62-112/Domain: myb DNA-binding repeat homology <MYB2>

Query Match	27.3%	Score 403;	DB 2;	Length 257;
Best Local Similarity	37.3%	Pred. No. 2.e-22;		
Matches	91;	Conservative 41;	Mismatches 68;	Indels 44;
				Gaps 7;
QY	11	VAKSGMTBEDDEFLTRKCIDKYGEGKMHLPVIRAGLNCRSCRCRLRWLYLPRHKGDFE	70	
		: : : : :           : : : : :           : : : : :		
Db	12	MKKGAMTKEDDLVDYIRKHGECWRSLPRAGLORCGKSCRLRMWNYLPLDLKRGNFT	71	
QY	71	ODEVDLILRLHLKLLGRMSLLIAGRLPGRTANDVKNYNTNLLRLKNTTKYIVR-EKINKK	129	
		: : : : :           : : : : :           : : : : :		
Db	72	EEDDELITLHSLGKMSLLIAGRLPGRTDNEIKNYNTTHIKRLLSRGIDPNSHRLINE	131	
QY	130	CGEISTKLEIIRKPRRRKYSSTMKVNTNNVYLDEEHCK-EIISKEQTPDASMDNVDPW	188	
		: : : : :           : : : : :           : : : : :		
Db	132	-----SVSP-----SSLQDVVERIHLDFSGVPKPEPRER-----	163	
QY	189	WINLLENCND-----DIEDEEVINYIEKTLTSLHBEISPLNIGEGSMOOGQISH	241	
		: : : : :           : : : : :           : : : : :		
Db	164	-IGMVNNCSSGTTSEKDYGNEDWVNLIELSVGPSYRYESTRKVSVDSAE-----ST	216	
QY	242	ENWG 245		
Db	217	RRWG 220		

RESULT 11  
T09773  
myb-related protein - upland cotton  
N:Alternate names: MYB-like DNA-binding domain protein  
C:Species: *Gossypium hirsutum* (upland cotton)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 19-May-2000  
C:Accession: T09773  
R:Roguerccio, L.L., Zhang, J., Wilkins, T.A.  
submitted to the EMBL Data Library, November, 1997  
A:Description: Structure and expression of six classes of myb-domain genes in allotetrap  
A:Reference number: Z16842  
A:Accession: T09773  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-226 <LOG>  
A:Cross-references: EMBL:AF034134, NID:g2921339, PID:g2921340

A:Experimental source: cultivar Acala Su-2; ovule  
 C/Genetics:  
 A/Note: Cmy-O  
 C/Superfamily: unassigned myb DNA-binding repeat proteins; myb DNA-binding repeat homolo  
 C/Keywords: DNA binding; transcription regulation  
 F:8-60/Domain: myb DNA-binding repeat homology <MYB>

Query Match 27.3%; Score 402.5; DB 2; Length 226;  
 Best Local Similarity 41.7%; Pred. No. 2.1e-22;  
 Matches 95; Conservative 34; Mismatches 66; Indels 33; Gaps 8;

QY 10 GVRKGSWTDEEDFLRKCIDKYGEKMHLPVIRAGLNRCRCKSLRLMNTLRPHIKRGDPF 69  
 DB 10 GTRKGAWSKQEDOKLIDYIRIHGECWRSLEKAGLHRCGSCSLRLMNTLRPHIKRGNF 69  
 QY 70 EODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLRLKLTNTKIVPREKINNK 129  
 DB 70 AODEEDLILRLHKLGNRWSLIAGRLPGRTDNEKYNWNSHIRKLMKMGIDPN--NHK 126  
 QY 130 CGEISTKIEIKPQRRKYFSSTMK---NVTNNVILDEEHCKEIISEKOTPDASMDNV 185  
 DB 127 LQYPHHVGPLNPTTNSMDVACSLRVCTDNDGSDASYL-----EDATPTGTISNL 181  
 QY 186 DPMWNI-----LNCNDIIEDEE---VVINTK--TITSLL 218  
 DB 182 D---LDLTIAFPSSPIKNI---IESOOKTASIVTNDDEEQYVPTLL 223

## RESULT 12

T51420  
 myb transcription factor werewolf WER/MYB66 - Arabidopsis thaliana  
 N/Alternate names: protein T9L3 50  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 02-Sep-2000  
 C/Accession: T51420  
 R/Seto, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew  
 submitted to the Protein Sequence Database, August 2000  
 A/Reference number: Z25394  
 A/Accession: T51420  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-203 <SAT>  
 A/Cross-references: EMBL:ALJ91149  
 A/Experimental source: cultivar Columbia; BAC clone T9L3  
 C/Genetics:  
 A/Map position: 5  
 A/Intons: 49/1; 92/2  
 A/Note: T9L3 50  
 C/Superfamily: trichome differentiation protein GLI; myb DNA-binding repeat homology

Query Match 27.1%; Score 401; DB 2; Length 203;  
 Best Local Similarity 48.7%; Pred. No. 2.3e-22;  
 Matches 77; Conservative 26; Mismatches 49; Indels 6; Gaps 2;

QY 12 RKGSWTDEEDFLRKCIDKYGEKMHLPVIRAGLNRCRCKSLRLMNTLRPHIKRGDPF 71  
 DB 17 KKGMTVEEDKILMDYKAKHGKHMRIAKKTGLKRGCKSLRLMNTLRPHIKRGDPF 76  
 QY 72 DEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLRLKLTNTKIVPREKINNK 131  
 DB 77 QOEEDLILRLHKLGNRWSLIAGRLPGRTDNEKYNWNSHIRKLMKMGIDPN--NHK 126  
 QY 132 EISTKIEIKPQRRKYFSSTMKNTNNVILDE--EEH 167  
 DB 133 DIVQINLPMTSETSEETKISNIVDNNITLGDIEQEDH 170

## RESULT 13

T09744  
 myb-related protein - upland cotton  
 N/Alternate names: MYB-like DNA-binding domain protein  
 C/Species: Gossypium hirsutum (upland cotton)  
 C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 19-May-2000

C/Accession: T09744  
 R/Loguercio, L.L.; Zhang, J.; Wilkins, T.A.  
 submitted to the EMBL Data Library, November 1997  
 A/Description: Structure and expression of six classes of myb-domain genes in allotetrapl  
 A/Reference number: Z16842

A/Accession: T09744  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-193 <LOG>  
 A/Cross-references: EMBL:AF034131; NID:92921333; PID:92921334  
 A/Experimental source: cultivar Acala Su-2; ovule  
 C/Genetics:  
 A/Gene: Cmy-G  
 C/Superfamily: unassigned myb DNA-binding repeat proteins; myb DNA-binding repeat homolo  
 C/Keywords: DNA binding; transcription regulation  
 F:62-112/Domain: myb DNA-binding repeat homology <MYB>

Query Match 27.1%; Score 400; DB 2; Length 193;  
 Best Local Similarity 43.3%; Pred. No. 2.6e-22;  
 Matches 88; Conservative 25; Mismatches 62; Indels 28; Gaps 5;

QY 5 SMSLSGVRKGSWTDEEDFLRKCIDKYGEKMHLPVIRAGLNRCRCKSLRLMNTLRPHI 64  
 DB 6 SLKRPVYKGAWTAEDRKLAEVITVHGAQRMWTTIPISAGLNRCGSKSLRLMNTLRPNI 65  
 QY 65 KRQDFODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLRLKLTNTKIVPRE 124  
 DB 66 KRGNISQDEEDLILRLHKLGNRWSLIAGRLPGRTDNEKYNWNSHIRKLMKMGIDPN--NHK 120  
 QY 125 KINNKGEISTK-----IEIKPQRRKYFSSTMKNTNNVILDEEHCKEIISEKOTP 178  
 DB 121 KHGASARQCKFAOQRLVNAKEQVEEWTST--GRGESNISFDVDF----- 167  
 QY 179 DASMDNVDF--WVINLNCND 198  
 DB 168 -FDPSNDVTNPFVWNRFLFVDD 189

## RESULT 14

T03972  
 anthocyanin biosynthesis regulatory protein Pl - maize  
 C/Species: Zea mays (maize)  
 C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 13-Aug-1999  
 C/Accession: T03972  
 R/Cone, K.C.; Cocciglione, S.M.; Moehlenkamp, C.A.; Weber, T.; Drummond, B.J.; Tagliani, I  
 Plant Cell 5, 1807-1816, 1993  
 A/Title: Role of the regulatory gene pl in the photocontrol of maize anthocyanin pigment  
 A/Reference number: Z15034; MUID:94138244; PMID:8305873  
 A/Accession: T03972  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-266 <CON>  
 A/Cross-references: EMBL:L19495; NID:9309569; PID:AAA19821.1; PID:9309570  
 A/Experimental source: strain McClinlock  
 C/Genetics:  
 A/Gene: pl  
 A/Intons: 45/1; 88/2  
 A/Function:  
 C/Description: transcriptional activator for anthocyanin biosynthesis  
 C/Superfamily: anthocyanin biosynthesis regulatory protein; myb DNA-binding repeat homolo  
 F:9-61/Domain: myb DNA-binding repeat homology <MYB>  
 F:62-112/Domain: myb DNA-binding repeat homology <MYB1>

Query Match 26.9%; Score 397; DB 2; Length 266;  
 Best Local Similarity 67.6%; Pred. No. 6.4e-22;  
 Matches 71; Conservative 14; Mismatches 20; Indels 0; Gaps 0;

QY 10 GVRKGSWTDEEDFLRKCIDKYGEKMHLPVIRAGLNRCRCKSLRLMNTLRPHIKRGDPF 69  
 DB 11 GVRKGAWTAKEDDTLAAVYAHGSKREVYQKAGLRCKSLRLMNTLRPHIKRGNT 70  
 QY 70 EODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLRLKLTNTKIVPREKINNK 114



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 20:04:26 ; Search time 18 Seconds  
(without alignments)  
715.851 Million cell updates/sec

Title: US-10-033-190-2  
Perfect score: 1477  
Sequence: 1 MNSTMSLSLGVKSGWTDEE.....QQGVQNDPSAIDLMNLID 274

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	411.5	27.9	258	1	TT2_ARATH
2	395	26.7	252	1	MYB1_MAIZE
3	394	26.7	273	1	MYB1_MAIZE
4	378	25.6	267	1	MYB1_HORVU
5	365	24.7	340	1	MYB1_MAIZE
6	362	24.6	255	1	MYB1_MAIZE
7	362	24.5	339	1	MYB1_MAIZE
8	361	24.4	228	1	GL1_ARATH
9	343	23.2	421	1	MYB2_PHYPA
10	339.5	23.0	302	1	MYB3_HORVU
11	263	17.8	743	1	MYB3_XENLA
12	262	17.7	704	1	MYB3_MOUSE
13	261.5	17.7	700	1	MYB3_MOUSE
14	255	17.3	686	1	MYB3_CHICK
15	252	17.1	657	1	MYB3_CHICK
16	248.5	16.8	728	1	MYB4_XENLA
17	242.5	16.4	751	1	MYB4_XENLA
18	242.5	16.4	752	1	MYB4_XENLA
19	241	16.3	624	1	MYB_XENLA
20	240.5	16.3	640	1	MYB_HUMAN
21	237.5	16.1	757	1	MYB_HUMAN
22	236.5	16.0	636	1	MYB_MOUSE
23	236.5	16.0	640	1	MYB_MOUSE
24	236.5	16.0	641	1	MYB_CHICK
25	235.5	15.9	382	1	MYB_AVIMB
26	235	15.9	367	1	ASL1_ARATH
27	221.5	15.0	451	1	MYB1_DICDI
28	185.5	12.6	429	1	MYB1_NEUCR
29	167	11.3	757	1	CC5_SCHPO
30	142.5	9.6	811	1	BAS1_YEAST
31	137	9.3	590	1	CBF1_YEAST
32	127.5	8.6	449	1	MYB1_PHYPA
33	108	7.3	595	1	REB1_XLUDA

34	101.5	6.9	402	1	ODP2_BUCAP	Q8K9E8 buchiera ap
35	100	6.8	810	1	REB1_YEAST	P21538 saccharomyc
36	99	6.7	639	1	TET2_ENTFA	P21598 enterococcu
37	99	6.7	639	1	TETM_STPAU	O53770 streptococc
38	99	6.7	639	1	TETM_STPAU	Q54807 streptococc
39	98.5	6.7	853	1	YCG1_YEAST	P25588 saccharomyc
40	98	6.6	639	1	TET5_ENTFA	P11131 enterococcu
41	98	6.6	639	1	TETM_NEIME	Q51238 nisseria m
42	97	6.6	608	1	DRTS_PLAFK	P13922 plasmodium
43	95.5	6.5	400	1	OMPA_THEMEA	O01969 thermotoga
44	95.5	6.5	510	1	PUSH_ARATH	O22928 arabidopsis
45	92.5	6.3	981	1	GLNE_HABIN	P44119 haemophilus

## ALIGNMENTS

RESULT 1  
TT2\_ARATH STANDARD; PRT; 258 AA.  
AC QPFFA2;  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE TRANSPARENT TESTA 2 protein (Myb-related protein 123) (AtMYB123) (Myb-  
related transcription factor IBM2-like).  
GN TT2 OR MYB123 OR ATMG35550 OR MOK9.18.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
[1]  
RP SEQUENCE FROM N.A., AND MUTANTS TT2-2 AND TT2-4.  
RC STRAIN=cv. Columbia, cv. Landsberg erecta, and cv. Wassilewskija;  
RX MEDLINE=21434422; PubMed=11549766;  
RA Neel N., Jond C., Debeaujon I., Caboche M., Lepoint L.;  
RT "The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts  
as a key determinant for proanthocyanidin accumulation in developing  
seed.";  
RL Plant Cell 13:2099-2114(2001).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=21481677; PubMed=11597504;  
RA Stracke R., Weider W., Weisshaar B.;  
RT "The R2R3-MYB gene family in Arabidopsis thaliana.";  
RL Curr. Opin. Plant Biol. 4:447-456(2001).  
[3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=99087489; PubMed=9872454;  
RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,  
Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.  
Sequence features of the regions of 1,013,767 bp covered by sixteen  
physically assigned P1 and TAC clones.";  
RL DNA Res. 5:297-308(1998).  
-1- FUNCTION: Transcription factor involved in the control of  
flavonoid late metabolism in developing siliques. Plays a key role  
in determining the tissue-specific activation of  
leucoanthocyanidin reductase (BAN1/ULS).  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- TISSUE SPECIFICITY: Expressed at a high level in immature siliques  
and at a lower level in flowers. Undetected in young seedlings,  
roots, leaves and inflorescence stems.  
CC -1- DEVELOPMENTAL STAGE: Highly expressed from the very early stages  
of embryogenesis to the globular stage, decreases rapidly from  
the late heart-shaped stage and did not persist after the  
completion of embryogenesis.  
CC -1- MISCELLANEOUS: TT2 activity is tightly linked to the presence of  
TT8.  
CC -1- SIMILARITY: Contains 2 Myb-like domains.



```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ299452; CAC00021.1; -
DR EMBL; AF371981; AAK54744.1; -
DR EMBL; AB015477; BAB08716.1; -
DR HSSP; P06876; IMBX.
DR InterPro: IPR001005; Myb DNA binding.
DR Pfam; PF00249; myb DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS50090; MYB_3; 2.
KW Flavonoid biosynthesis; Transcription; Transcription regulation;
KW Nuclear protein; ATP-binding; DNA-binding; Repeat; Polymorphism.
FT DNA_BIND 11 63 MYB 1.
FT DNA_BIND 64 114 MYB 2.
FT NP_BIND 47 54 ATP (POTENTIAL).
FT VARIANT 174 174 Q -> L (IN CV. WS).
FT MUTAGEN 66 66 G->R; IN TT2-4; LOSS OF SEED
FT MUTAGEN 127 258 PIGMENTATION.
FT MUTAGEN 127 258 MISSING; IN TT2-2; REDUCED SEED
FT SEQUENCE 258 AA; 29611 MW; 3BA9CB3F723D1C2E CRC64;
SQ
Query Match 27.9%; Score 411.5; DB 1; Length 258;
Best Local Similarity 42.9%; Pred. No. 2.9e-23;
Matches 96; Conservative 32; Mismatches 65; Indels 27; Gaps 6;
QY 3 STNSSLGVRKGSWTDEDFLRKCIDKYGGKMLVPIRAGLNCRCKSCLRLMYLRLP 62
DB 6 TTSVRBELNKGAMTWDHEDKILRDYITTHGEGKSTLPNGGLKRCGKSCRLRMNYLRP 65
QY 63 HIRKGDPEQDEVDLIRLHLKLNRMWSLIAGRLPGRTPANDVKNYNTNLRLKNTTKIVP 122
DB 66 GIKKGNISSEDEELIRLHLNLLGNRWSLIAGRLPGRTPANEIKMNSMLRRLPRTQKQ 125
QY 123 REKI-----NNKGEISTKLEIKPQRKYFS--STMKNVNTNNVILDEE----- 166
DB 126 PKRIKHSNNENNVN--VIRTK--AIRCKTLLFSLSLQKSSSTSPPLPKQENDQGGSS 182
QY 167 -----HCKEIISEKQTPD--ASMDNVDPWMVINLIENCNDIEE 202
DB 183 LMGDLFPDFRIHSEFFPDLMDFDGLDCGNVTSIVSSNEITLGE 226
RESULT 2
MYBD_MAIZE STANDARD; PRT; 252 AA.
ID MYBD_MAIZE
AC P23592;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anthocyanin regulatory Cl-I protein.
GN Cl.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90151603; PubMed=2303027;
RA Paz-Ares J., Ghosal D., Saedler H.;
RT "Molecular analysis of the Cl-I allele from Zea mays: a dominant
RT mutant of the regulatory Cl locus.";
RL EMBO J. 9:315-321(1990).

```

```

CC -----
CC -1- FUNCTION: CONTROLS THE EXPRESSION OF GENES INVOLVED IN ANTHOCYANIN
CC BIOSYNTHESIS. REGULATES THE EXPRESSION OF AT LEAST 3 STRUCTURAL
CC GENES: CHALCONE SYNTHASE, DIHYDROFLAVONOL REDUCTASE AND FLAVONOL
CC O(3) GLUCOSYLTRANSFERASE. Cl ACTS AS A TRANS-ACTING FACTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable)
CC -1- MISCELLANEOUS: THE Cl-1 ALLELE IS A DOMINANT NEGATIVE MUTANT WHICH
CC INHIBITS PIGMENT FORMATION.
CC -1- SIMILARITY: Contains 2 Myb-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52201; CA36456.1; -
DR HSSP; P01103; IPOW.
DR TRANSFAC; T02946; -.
DR MaizeDB; 24964; -.
DR InterPro: IPR001005; Myb DNA binding.
DR Pfam; PF00249; myb DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS50090; MYB_3; 2.
KW Nuclear protein; DNA-binding; Repeat; Transcription regulation;
KW Activator; Trans-acting factor.
FT DNA_BIND 9 61 MYB 1.
FT DNA_BIND 62 112 MYB 2.
FT DOMAIN 203 212 POLY-GLY.
FT SEQUENCE 252 AA; 26419 MW; AD61FA2B829C5FBD CRC64;
SQ
Query Match 26.7%; Score 395; DB 1; Length 252;
Best Local Similarity 65.7%; Pred. No. 4.4e-22;
Matches 69; Conservative 16; Mismatches 20; Indels 0; Gaps 0;
QY 10 GVRKGSWTDEDFLRKCIDKYGGKMLVPIRAGLNCRCKSCLRLMYLRLPPIKSGDF 69
DB 11 GVRGANTSTSEDALAVYVAHSGKREYVQKAGLRCKSCRLRLMYLRLPIRIRGNI 70
QY 70 EODEVDLIRLHLKLNRMWSLIAGRLPGRTPANDVKNYNTNLRLK 114
DB 71 SYDEEDLIRLHLNLLGNRWSLIAGRLPGRTPANEIKMNSMLRRLPRTQKQ 115
RESULT 3
MYBC_MAIZE STANDARD; PRT; 273 AA.
ID MYBC_MAIZE
AC P10290;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anthocyanin regulatory Cl-I protein.
GN Cl.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88111545; PubMed=3428265;
RA Paz-Ares J., Ghosal D., Wienand U., Peterson P.A., Saedler H.;
RT "The regulatory Cl locus of Zea mays encodes a protein with homology
RT to myb proto-oncogene products and with structural similarities to
RT transcriptional activators.";
RL EMBO J. 6:3553-3558(1987).
-1- FUNCTION: CONTROLS THE EXPRESSION OF GENES INVOLVED IN ANTHOCYANIN
CC BIOSYNTHESIS. REGULATES THE EXPRESSION OF AT LEAST 3 STRUCTURAL
CC GENES: CHALCONE SYNTHASE, DIHYDROFLAVONOL REDUCTASE AND FLAVONOL
CC O(3) GLUCOSYLTRANSFERASE. Cl ACTS AS A TRANS-ACTING FACTOR.

```



```

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 2 Myb-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M37153, AAA33482.1, -.
DR PIR, S06215, TVZAMB.
DR HSSP, P01103, 1POM.
DR TRANSPAC, T01592, -.
DR MaizeDB, 24964, -.
DR InterPro, IPR001005, Myb DNA-binding.
DR Pfam, PF00249, myb DNA-binding; 2.
DR SMART, SM00717, SANT; 2.
DR PROSITE, PS00037, MYB_1; 1.
DR PROSITE, PS00334, MYB_2; 1.
DR PROSITE, PS50090, MYB_3; 2.
DR Nuclear protein; DNA-binding; Repeat; Transcription regulation;
DR Activator; Trans-acting factor.
DR DNA_BIND 9 61 MYB 1.
FT DNA_BIND 62 112 MYB 2.
FT DOMAIN 205 212 POLY-GLY.
FT DOMAIN 234 273 ASF/GLU-RICH (ACIDIC).
SQ SEQUENCE 273 AA; 28750 MW; 163977BBA8CE669D CRC64;

Query Match 26.7%; Score 394; DB 1; Length 273;
Best Local Similarity 65.7%; Pred. No. 5.8e-22;
Matches 69; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

QY 10 GVRKGSWTDEEDFLRKCDIKYGEKWLVPVIRAGLNRCKRSCLRLWNLIRPHIKGDF 69
DB 11 GVKRGAWTSKEDDLAAVYKKAHGEGKWEVPOKAGLRCKGSCRLRWNLIRPNIIRGNI 70
QY 70 EODVDLIRLRLKLGKRWMSLIGRLPGRTANDVKNYWNNTLAK 114
DB 71 SYDEDLIRLRLKLGKRWMSLIGRLPGRTDNEIKYWNSTLGR 115

RESULT 4
MYB1 HORVU STANDARD; PRT; 267 AA.
AC P20026;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myb-related protein Hv1.
GN MYB1.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Hordeum.
OC NCB1_TaxID=4513;
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Abyssinian 2231; PubMed=2664447;
RX MEDLINE=89313655;
RA Marocco A., Wiesenbach M., Becker D., Paz-Ares J., Saedler H.,
RA Salamini F., Rohde W.;
RT "Multiple genes are transcribed in Hordeum vulgare and Zea mays that
RT carry the DNA binding domain of the myb oncoproteins.";
RL Mol. Gen. Genet. 216:183-187(1989).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Abyssinian 2231; TISSUE=Leaf;
RX MEDLINE=94035190; PubMed=8220488;
RA Wiesenbach M., Ueberlackner B., Vogt F., Becker D., Salamini F.,
RA Rohde W.;
RT "Myb genes from Hordeum vulgare: tissue-specific expression of

```

```

RT chimeric Myb promoter/Gus genes in transgenic tobacco.";
RL Plant J. 4:411-422(1993).
CC -1- FUNCTION: POSSIBLE TRANSCRIPTION ACTIVATOR IN RESPONSE TO AN
CC EXTERNAL SIGNAL. MAY BE INVOLVED IN THE REGULATION OF FLAVONOID
CC BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: GERMINATING SEED AND APICAL MERISTEM OF SHOOT
CC AND ROOT.
CC -----
CC -1- SIMILARITY: Contains 2 Myb-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X70877, CA50222.1, -.
DR EMBL, X70879, CA50224.1, -.
DR PIR, S61506, S61506.
DR HSSP, P01103, 1POM.
DR TRANSPAC, T02887, -.
DR InterPro, IPR001005, Myb DNA-binding.
DR Pfam, PF00249, myb DNA-binding; 2.
DR SMART, SM00717, SANT; 2.
DR PROSITE, PS00037, MYB_1; 1.
DR PROSITE, PS00334, MYB_2; 1.
DR PROSITE, PS50090, MYB_3; 2.
DR Nuclear protein; DNA-binding; Activator; Transcription regulation;
DR Repeat.
DR DNA_BIND 9 61 MYB 1.
FT DNA_BIND 62 112 MYB 2.
SQ SEQUENCE 267 AA; 29740 MW; DBECDCEBA847749 CRC64;

Query Match 25.6%; Score 378; DB 1; Length 267;
Best Local Similarity 54.5%; Pred. No. 8.3e-21;
Matches 73; Conservative 18; Mismatches 41; Indels 2; Gaps 1;

QY 13 KGSWTDEEDFLRKCDIKYGEKWLVPVIRAGLNRCKRSCLRLWNLIRPHIKGDF 72
DB 14 KGAWTKEEDRLTAIYAHGEGCWRSLPKAAGLRCKGSCRLRWNLIRPDLKGNFSHE 73
QY 73 EVDLIRLRLKLGKRWMSLIGRLPGRTANDVKNYWNNTLAKNTTIVP--REKINNC 130
DB 74 EDELIRKLSLGLGKRWMSLIGRLPGRTDNEIKYWNSTLGR 133
QY 131 GEISTKLEIKPOR 144
DB 134 AASWITSPESQOR 147

RESULT 5
MYB1 MAIZE STANDARD; PRT; 340 AA.
AC P20024;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myb-related protein Zm1.
GN Zea mays (Maize).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OC NCB1_TaxID=4577;
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Abyssinian 2231; PubMed=2664447;
RX MEDLINE=89313655;
RA Marocco A., Wiesenbach M., Becker D., Paz-Ares J., Saedler H.,
RA Salamini F., Rohde W.;
RT "Multiple genes are transcribed in Hordeum vulgare and Zea mays that
RT carry the DNA binding domain of the myb oncoproteins.";
RL Mol. Gen. Genet. 216:183-187(1989).

```

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 2 Myb-like domains.  
 DR PIR; S04898; S04898.  
 DR HSP; P06876; IMK.  
 DR TRANSFAC; T02958; -.  
 DR MaizeDB; 69590; -.  
 DR InterPro; IPR001005; Myb DNA binding.  
 DR Pfam; PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00337; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KM Nuclear protein; DNA-binding; Repeat; Transcription regulation.  
 FT DNA\_BIND 11 63 MYB 1.  
 FT DNA\_BIND 64 114 MYB 2.  
 SQ SEQUENCE 340 AA; 36239 MW; F2341387F6263DAF CRC64;  
 Query Match 24.7%; Score 365; DB 1; Length 340;  
 Best Local Similarity 32.1%; Pred. No. 9.9e-20;  
 Matches 94; Conservative 45; Mismatches 106; Indels 48; Gaps 8;  
 QY 7 SSLGVRKGSWTDEDFLRKCIDKYEGSKHLVPIRAGLNRCRKSCLRWNLRIHRIKRGDEPD 66  
 DB 10 AKVGLNGSWTPQEDMLIAVIOKGHITNMRALPKQAGILRCGKSCRLRWNLRIHRIKRGDEPD 69  
 QY 67 GDFEODEVDLIRLHLKLNKWSLIGRLPGRTANDVKNYNTNLRLKLTNTKTVPEKI 126  
 DB 70 GNFTDEEBEALIRLHGLNGWSKIAACLPGRTDNEIKYNTNTHLKK-----KVAQHEKK 124  
 QY 127 NNKGEISTKIEIKPQRKFFSSTMKVNTNNVYLDEEHCKELISEKQTPDAS----- 181  
 DB 125 KAGGSGDAGCPATAPLSSATSTTTNNSSGSS---DSGDCQ-----GTSRRPDATDCTL 177  
 QY 182 ----MDNVD-----PWNILNLCNDIDEEDEY-VINYEKTLTSLHBE 221  
 DB 178 QPEDMDVSDMLVDGAPPAQPMPSRSSSSSLTTCVGVEELIELPVIDIPEEIMSIIDGE 237  
 QY 222 ISPLNTGEGNSMQCGQISHENWGFSLNLPMMQGVND-----DFSIEDLW 270  
 DB 238 SAVARHGDAAPCTGTGT-----AVSTSEAEBAANDWLEKEKELGLW 283

RESULT 6  
 MYB3\_MAIZE  
 ID MYB3\_MAIZE STANDARD; PRT; 255 AA.  
 AC P20025;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Myb-related protein Zm38.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89313655; PubMed=2664447;  
 RA Marocco A., Wiesenbach M., Becker D., Paz-Ares J., Seedler H.,  
 RA Salamini F., Rohde W.;  
 RT "Multiple genes are transcribed in Hordeum vulgare and Zea mays that  
 RT carry the DNA binding domain of the myb oncoproteins";  
 RL Mol. Gen. Genet. 216:183-187(1989).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 2 Myb-like domains.  
 DR HSP; P01103; IPOW.  
 DR TRANSFAC; T02959; -.  
 DR MaizeDB; 69594; -.  
 DR InterPro; IPR001005; Myb DNA binding.  
 DR Pfam; PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00337; MYB\_1; 1.

DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KM Nuclear protein; DNA-binding; Repeat; Transcription regulation.  
 FT DNA\_BIND 9 61 MYB 1.  
 FT DNA\_BIND 62 112 MYB 2.  
 SQ SEQUENCE 255 AA; 27568 MW; 26AADC8B0633E06 CRC64;  
 Query Match 24.6%; Score 363; DB 1; Length 255;  
 Best Local Similarity 64.1%; Pred. No. 9.7e-20;  
 Matches 66; Conservative 14; Mismatches 23; Indels 0; Gaps 0;  
 QY 13 KGSWTDEDFLRKCIDKYEGSKHLVPIRAGLNRCRKSCLRWNLRIHRIKRGDEPD 72  
 DB 14 KGATKEDERLVAIVIRAHGEGCWRSLPKAAGILRCGKSCRLRWNLRIHRIKRGDEPD 73  
 QY 73 EVDLIRLHLKLNKWSLIGRLPGRTANDVKNYNTNLRLKLTNTKTVPEKI 115  
 DB 74 EDDLIVRLHSLGNKWSLIAARLPGRTDNEIKYNTNTHLKK-----KVAQHEKK 116

RESULT 7  
 MYB3\_MAIZE  
 ID MYB3\_MAIZE STANDARD; PRT; 399 AA.  
 AC P27898; P27899;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Myb-related protein P.  
 GN P.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91271238; PubMed=2052542;  
 RA Grotenwald B., Althma P., Peterson T.;  
 RT "Alternatively spliced products of the maize P gene encode proteins  
 RT with homology to the DNA-binding domain of myb-like transcription  
 RT factors";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4587-4591(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92275319; PubMed=1317315;  
 RA Althma P., Grotenwald B., Peterson T.;  
 RT "Insertional mutagenesis of the maize P gene by intragenic  
 RT transposition of Ac.";  
 RL Genetics 131:199-209(1992).  
 CC -1- FUNCTION: TRANSCRIPTION FACTOR POSTULATED TO REGULATE THE  
 CC BIOSYNTHETIC PATHWAY OF A FLAVONOID-DERIVED PIGMENT IN CERTAIN  
 CC FLORAL TISSUES.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=long;  
 CC IsoId=P27898-1; Sequence=Displayed;  
 CC Name=short;  
 CC IsoId=P27898-2; Sequence=VSP\_003301, VSP\_003302;  
 CC -1- SIMILARITY: Contains 2 Myb-like domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

EMBL; M73028; AAA33500.1; -  
 DR EMBL; M73029; AAA33501.1; -  
 DR EMBL; Z11879; CAA77939.1; -  
 DR PIR; A39697; A39697.

```

DR PIR; B39697; B39697.
DR HSP; P01103; IPOM.
DR TRANSFAC; T01590; -.
DR TRANSFAC; T01591; -.
DR MaizeDB; 69180; -.
DR MaizeDB; 69181; -.
DR InterPro; IPR01005; Myb_DNA_binding.
DR Pfam; PF00249; myb_DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS50090; MYB_3; 2.
DR Nuclear protein; DNA-binding; Repeat; Transcription regulation;
KW Alternative splicing.
FT DNA_BIND 9 61 MYB 1.
FT DNA_BIND 62 112 MYB 2.
FT DOMAIN 178 181 POLY-ALA.
FT DOMAIN 305 311 POLY-LEU.
FT VARSPIC 89 152 WSLASHLPGRTDNEIKYNNWSHLRQIHTYRKYTAGPPD
TALAIDMSKLSADRRGGRTPG -> RHIMTEADYSPPT
VRCPLRGALAYLTLPKPSPTARITYDRIGSALLRSVRFC
PRCPSPRW (in isoform short).
FT FTID=VSP_003301.
FT VARSPIC 153 399 Missing (in isoform short).
FT FTID=VSP_003302.
SQ SEQUENCE 399 AA; 43756 MW; EE025B0D44CF5D0 CRC64;

Query Match 24.4%; Score 362; DB 1; Length 399;
Best Local Similarity 56.9%; Pred. No. 2e-19;
Matches 62; Conservative 22; Mismatches 25; Indels 0; Gaps 0;

QY 9 LGVKGSGWTDEEDFLRKCDIKYEGKWHLPVIRAGLNCRKSGRLRLMYLRIHRIKGD 68
DB 10 VGLKRGKWTAEEDDLLNYIAHEGSGRSIPKAGLLRCKSKRLKINILRADVKGK 69
QY 69 FEODEVDLIRLHKLGNNRSLIAGRLPGRTADVKNYNNNTLLRKLT 117
DB 70 ISKEEDILIKLHNTLGNRWSLIASHLPGRTDNEIKYNNWSHLRQIHT 118

RESULT 8
GL1 ARATH
ID -GL1 ARATH STANDARD; PRT; 228 AA.
AC P27500; O9ST42;
DT 01-AUG-1992 (Rel. 23, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Trichome differentiation protein GL1 (GLABROUS1 protein).
GN GL1 OR AT3G27920 OR K16N12.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92034971; PubMed=1934056;
RA Oppenheimer D.G., Herman P.L., Sivakumaran S., Bech J., Marks M.D.;
RT "A myb gene required for leaf trichome differentiation in Arabidopsis
is expressed in stipules";
RL Cell 67:483-493 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=99449058; PubMed=10520748;
RA Shikazono N., Tanaka A., Yokota Y., Watanabe H., Tano S.;
RT "Nucleotide sequence of the GLABROUS1 gene of Arabidopsis thaliana
ecotype Columbia.";
RL DNA Seq. 9:177-181 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia, cv. Ba-1, cv. Bla-1, cv. Cond, cv. Sha, and
RC cv. Te-0;

```

```

RX MEDLINE=21396402; PubMed=11504855;
RA Hauser M.T., Harr B., Schotterer C.;
RT "Trichome distribution in Arabidopsis thaliana and its close relative
Arabidopsis lyrata: molecular analysis of the candidate gene
GLABROUS1.";
RL Mol. Biol. Evol. 18:1754-1763 (2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 pl. TAC
and BAC clones.";
RL DNA Res. 7:217-221 (2000).
CC -1- FUNCTION. REGULATES THE PRODUCTION OF A SIGNAL THAT INDUCES
HAIR (TRICHOME) PRECURSOR CELLS ON LEAF PRIMORDIA TO
DIFFERENTIATE.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 2 Myb-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements
or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M79448; AAC97387.1; -.
DR EMBL; AB006078; BAA86879.1; -.
DR EMBL; AF263690; AAL01215.1; -.
DR EMBL; AF263693; AAL01218.1; -.
DR EMBL; AF263694; AAL01219.1; -.
DR EMBL; AF263695; AAL01220.1; -.
DR EMBL; AF263696; AAL01221.1; -.
DR EMBL; AF263697; AAL01222.1; -.
DR EMBL; AP000371; BAB02538.1; -.
DR HSP; P01103; IPOM.
DR TRANSFAC; T01588; -.
DR InterPro; IPR01005; Myb_DNA_binding.
DR Pfam; PF00249; myb_DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS50090; MYB_3; 2.
KW Nuclear protein; DNA-binding; Repeat; Transcription regulation.
FT DNA_BIND 11 63 MYB 1.
FT DNA_BIND 64 114 MYB 2.
FT CONFLICT 149 149 C -> R (IN REF. 1).
FT CONFLICT 222 224 GDV -> SDI (IN REF. 1).
SQ SEQUENCE 228 AA; 26338 MW; 822A8646F446FB9 CRC64;

Query Match 24.4%; Score 361; DB 1; Length 228;
Best Local Similarity 35.2%; Pred. No. 1.2e-19;
Matches 80; Conservative 39; Mismatches 58; Indels 50; Gaps 6;

QY 12 RKSGWTDEEDFLRKCDIKYEGKWHLPVIRAGLNCRKSGRLRLMYLRIHRIKGD 71
DB 15 KKGILWVEEDNIIIMDYLANHGTQWNRIVAKTGKCGSKRLRWNYLSPNKGAFTE 74
QY 72 DEVDLIRLHKLGNNRSLIAGRLPGRTADVKNYNNNTLLRKLT 121
DB 75 QEBDLIRLHKLGNNRSLIAGRLPGRTDVQVKNYNNHTLSKTL 120
QY 132 EISTKIEI-----IKPRKRYFSSTMKVNTNNVILDEEHCKEI 171
DB 121 DYSAAVTGTGEDDPSPLFITATPSSCHQDENIYENAKF-NGVVSASVDEKPKOE 179
QY 172 ISEKQTPDASMDNDP-----WVNIENLNCNDIIEDEEVINY 210
DB 180 LAQKDVLMATYTN--DPSHYYGNNALWVH-----DDDFEISLVMMNF 219

```

```

RESULT 9
MYB2_PHYPA STANDARD: PRT: 421 AA.
ID MYB2_PHYPA
AC P80073;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myb-related protein Pp2.
GN Pp2.
OS Physcomitrella patens (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryophyta; Funariidae; Funariales; Funariaceae; Physcomitrella.
ON NCBI_Taxid=3218;
RX SEQUENCE FROM N.A.
MEDLINE=94004988; PubMed=8401607;
RA Leech M.J., Kammerer W., Cove D.J., Martin C., Wang T.L.;
RT "Expression of myb-related genes in the moss, Physcomitrella patens.";
RL Plant J. 3:51-61(1993).
CC -1- FUNCTION: POSSIBLE TRANSCRIPTION ACTIVATOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- DEVELOPMENTAL STAGE: HIGH RATES OF GROWTH.
CC -1- SIMILARITY: Contains 2 Myb-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X67050; CAA47435.1; -.
CC PIR, S24244; S24244.
CC HSSP, P01103; 1POM.
CC TRANSFAC; T02893; -.
CC InterPro; IPR001005; Myb DNA binding.
CC Pfam; PF00249; myb DNA-binding; 2.
CC SMART; SM00717; SANT; 2.
CC PROSITE; PS00037; MYB_1; 1.
CC PROSITE; PS00334; MYB_2; 1.
CC PROSITE; PS00090; MYB_3; 2.
CC KW Nuclear protein; DNA-binding; Repeat; Transcription regulation;
CC Activator.
CC FT DNA_BIND 61 MYB 1.
CC FT DNA_BIND 62 112 MYB 2.
CC SQ SEQUENCE 421 AA; 46695 MW; 27A3578578D0F3F4 CRC64;
Query Match 23.2%; Score 343; DB 1; Length 421;
Best Local Similarity 34.3%; Pred. No. 5.2e-18;
Matches 80; Conservative 34; Mismatches 95; Indels 24; Gaps 3;
QY 9 LGVRKSWTDEEDFLRKCDIKYGEKWHLPVRAGLNRCRKSCLRLMYLNRPHIKGD 68
DB 10 VGLRGPMTSEBDQVLGHITNNGLSCRAIPKLAGLRCCSGCLRTMYTLRLPDLKKGI 69
QY 69 FEODEVDLILRLHKLGNRWSLIAGRLPGRANDVYKATWNTNLAKNTTKIVPREKLN 128
DB 70 FSEAENILDLHATLGNRWSRIAQLPGRTDNEIKYNTWNTLKKRLSSQGLDPVTHLP 129
QY 129 KGEISTEIEIKPKRRKYFSSITKGNVNNVILDEEBECKEIISEKO-----TPDASM 182
DB 130 EDSKLDJTEDTDDGSSDYNTMSDASKSEKSKKKPKETVAVRQPKGPKAPQLKM 189
QY 183 DNVD-----PWWINLLENCNDIDEEDEVIVYKTLTSLHEE 221
DB 190 QSDSGPVLKPKAPKSPISVNPGRGNYDDDSH-----SSSSVTTTKSHED 238

```

```

AC P20027;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myb-related protein Hv33.
GN MYB2.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Hordeum.
ON NCBI_Taxid=4513;
RX SEQUENCE FROM N.A.
MEDLINE=94035190; PubMed=8220488;
RA Wissenbach M., Ueberlacker B., Vogt F., Becker D., Salamini F.,
RA Ronde W.;
RT "Myb genes from Hordeum vulgare: tissue-specific expression of
RT chimeric Myb promoter/Gus genes in transgenic tobacco.";
RL Plant J. 4:411-422(1993).
CC [2]
CC SEQUENCE OF 35-302 FROM N.A.
CC STRAIN=cv. Abyssinian 2231;
CC MEDLINE=93333655; PubMed=264447;
CC Marocco A., Wissenbach M., Becker D., Paz-Ares J., Saedler H.,
CC Salamini F., Ronde W.;
CC "Multiple genes are transcribed in Hordeum vulgare and Zea mays that
CC carry the DNA binding domain of the myb oncoproteins.";
CC Mol. Genet. 216:183-187(1998).
CC -1- FUNCTION: POSSIBLE TRANSCRIPTION ACTIVATOR IN RESPONSE TO AN
CC EXTERNAL SIGNAL. MAY BE INVOLVED IN THE REGULATION OF FLAVONOID
CC BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: GERMINATING SEED AND APICAL MERISTEM OF SHOOT
CC AND ROOT.
CC -1- SIMILARITY: Contains 2 Myb-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X70881; CAA50226.1; -.
CC PIR, X70878; CAA50223.1; -.
CC HSSP, P01103; 1POM.
CC TRANSFAC; T02889; -.
CC InterPro; IPR001005; Myb DNA binding.
CC Pfam; PF00249; myb DNA-binding; 2.
CC SMART; SM00717; SANT; 2.
CC PROSITE; PS00037; MYB_1; 1.
CC PROSITE; PS00334; MYB_2; 1.
CC PROSITE; PS00090; MYB_3; 2.
CC KW Nuclear protein; DNA-binding; Activator; Transcription regulation;
CC Repeat.
CC FT DNA_BIND 11 MYB 1.
CC FT DNA_BIND 64 114 MYB 2.
CC FT CONFLICT 259 A->P (IN REF. 2).
CC SQ SEQUENCE 302 AA; 33003 MW; 89017418C3631599 CRC64;
Query Match 23.0%; Score 339.5; DB 1; Length 302;
Best Local Similarity 53.6%; Pred. No. 6.2e-18;
Matches 67; Conservative 16; Mismatches 39; Indels 3; Gaps 1;
QY 1 MNSTSMSSLG---VRKSWTDEEDFLRKCDIKYGEKWHLPVRAGLNRCRKSCLRLRWL 57
DB 1 MGRPSSGAVGQPKYRKRLMSEEDBKLYNHIIRRGVCSVPRLLAALNRCGSCRLRWI 60
QY 58 NYLRPHIKRGDFEODEVDLILRLHKLGNRWSLIAGRLPGRANDVYKATWNTNLRLKLT 117

```

```

Db      61  NYLRFDLRGCPSSQGEEDHIVAHQIILGNRMSQIASHLPGRTDNEIKAFMNSCIKKLRLQ 120
Cy      118  TKIYP 122
Db      121  OGIDP 125

RESULT 11
MYBB_XENLA
ID MYBB_XENLA STANDARD; PRT; 743 AA.
AC P52551;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Myb-related protein B (B-Myb) (Myb-related protein 1) (XMYB1).
OS MYBL2 OR BMYB OR MYBL.
OC Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=83355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92297434; PubMed=1606020;
RA Bouwmeester T., Guenemann S., El-Barradi T., Kalbrenner F.,
RA van Wijk I., Moelling K., Pielier T.;
RT "Molecular cloning, expression and in vitro functional
RT characterization of Myb-related proteins in Xenopus.";
RL Mech. Dev. 37:57-68 (1992).
RL [2]
RN REVISTONS TO C-TERMINUS.
RX MEDLINE=99214591; PubMed=10187816;
RA Humbert-Lan G., Pielier T.;
RT "Regulation of DNA binding activity and nuclear transport of B-Myb in
RT Xenopus oocytes.";
RL J. Biol. Chem. 274:10293-10300 (1999).
RL -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- DEVELOPMENTAL STAGE: PRESENT THROUGHOUT OOGENESIS AND EARLY
CC XENOPOUS EMBRYOGENESIS; IN ADULT TISSUE IT IS PRIMARILY DETECTED IN
CC BLOOD.
CC -I- SIMILARITY: Contains 3 Myb-like domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC CC
CC EMBL; M75870; AAC98701.1; -.
CC DR HSSP; 003237; IASJ.
CC DR InterPro: IPR001005; Myb_DNA_binding.
CC DR Pfam; PF00249; Myb_DNA-binding; 3.
CC DR SMART; SM00717; SANT; 3.
CC DR PROSITE; PS00037; MYB 1; 3.
CC DR PROSITE; PS00334; MYB 2; 3.
CC DR PROSITE; PS00090; MYB 3; 3.
CC KW Transcription regulation; Nuclear protein; DNA-binding; Repeat.
FT DNA_BIND 26 77 MYB 1.
FT FT DNA_BIND 78 129 MYB 2.
FT FT DNA_BIND 130 180 MYB 3.
SQ SEQUENCE 743 AA; 82909 MW; 5FD1D678BB24409B CRC64;

Query Match 17.8%; Score 263; DB 1; Length 743;
Beet Local Similarity 31.9%; Pred. No. 7.2e-12;
Matches 69; Conservative 36; Mismatches 83; Indels 28; Gaps 6

13 KGSWTDEEDFLRKCTIDKXGEGKMLVPIRAGLNRCSRLRMVNYLRPHIKRGDFQD 72
Db 83 KGPWKEDEBKETELVKKYGTGHWLLI-AKQLRGSMGQCRRRMHNLNPVKSKSWTEE 141
Cy 73 EVDLILRLHLKLGNEWSLIAGRLPGRITANDVKNVNTLLRKLTNTKTV-----PR 123

```

[illegible]



DR SMART; SM00717; SANT; 3.  
DR PROSITE; PS00037; MYB 1; 3.  
DR PROSITE; PS00334; MYB 2; 3.  
DR PROSITE; PS50090; MYB 3; 3.  
KW Transcription regulation; Nuclear protein; DNA-binding; Repeat;  
Phosphorylation.  
FT DNA\_BIND 26 77 MYB 1.  
FT DNA\_BIND 78 129 MYB 2.  
FT DNA\_BIND 130 180 MYB 3.  
FT MOD\_RES 440 440 PHOSPHORYLATION (BY CDK2).  
FT MOD\_RES 444 444 PHOSPHORYLATION (BY CDK2).  
FT MOD\_RES 487 487 PHOSPHORYLATION (BY CDK2).  
FT MOD\_RES 494 494 PHOSPHORYLATION (BY CDK2).  
FT MOD\_RES 520 520 PHOSPHORYLATION (BY CDK2).  
FT MOD\_RES 577 577 PHOSPHORYLATION (BY CDK2).  
FT MOD\_RES 577 577 PHOSPHORYLATION (BY CDK2).  
SQ SEQUENCE 700 AA; 78764 MW; D91B28B3DAB94061 CRC64;  
Query Match 17.7%; Score 261.5; DB 1; Length 700;  
Best Local Similarity 30.1%; Pred. No. 8.6e-12;  
Matches 88; Conservative 40; Mismatches 101; Indels 63; Gaps 12;  
QY 13 KGSWTDEDFLRKCIDKYGKWHLPVIRAGLNRCKRSLRLMYLRPHIKRGDFEOD 72  
DB 83 KGPWTKEDQKVIKLVKYGKWHLPVIRAGLNRCKRSLRLMYLRPHIKRGDFEOD 141  
QY 73 EVDILRLHLKLGKRWSLIAGRLPGRTANDVKNYNTNLRLKNTTKIVPREKINCKG- 131  
DB 142 EDRITCEAHKVLGNRMALIAKMLPGRTDNAVKNHNSITKRVDTGGLNETK--DCKP 198  
QY 132 -----EISTKIEIKPQRRKYSSTMKVNTNNVILDEEHCKEII-----SEKQTP-D 179  
DB 199 PVYLLLEDEKQGLQSAQPTGQGLTLTNWPSVPTKEEENSEBELAAATTSKQEPFG 258  
QY 180 ASMNVD-----PM-WI-----NL-----LNCNDDEED 203  
DB 259 TDLDAVTPPELPEPKKEDEGSPETSPLPYKVVANLILPAVGSLSLEALDLISD 318  
QY 204 EEVYINYEKTLTSLHEEISPLNTIGEGNSMOQGOISHENGEFSLNLPWQ 255  
DB 319 PDACNDLAK--FDLPPEPSADSI--NNSLVQLQASHQ-----QVLLPQK 360  
RESULT 14  
MYB\_CHICK  
ID MYB\_CHICK STANDARD; PRT; 686 AA.  
AC Q00237;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE MYB-related protein B (B-Myb).  
GN MYB12 OR BMYB.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93049214; PubMed=1425593;  
RA Foos G., Grimm S., Klemmner K.-H.;  
RT "Functional antagonism between members of the myb family: B-myb  
inhibits v-myb-induced gene activation.";  
RL EMO J. 11:4619-4629(1992).  
RN [2]  
RP STRUCTURE BY NMR OF 79-186.  
RX MEDLINE=98322063; PubMed=9657674;  
RA McIntosh P.B., Frenkel T.A., Mollborn U., McCormick J.E.,  
RA Klemmner K.H., Feeney J., Carr M.D.;  
RT "Solution structure of the B-Myb DNA-binding domain: a possible role  
for conformational instability of the protein in DNA binding and  
control of gene expression."  
RL Biochemistry 37:9619-9629(1998).  
DE -I- FUNCTION: IT REPRESENTS V-MYB- AND C-MYB-MEDIATED ACTIVATION OF

CC THE MIM-1 GENE, PROBABLY BY COMPETING WITH OTHER MYB PROTEINS  
CC FOR BINDING SITES. IT IS AN INHIBITORY MEMBER OF THE MYB FAMILY.  
CC -I- SUBCELLULAR LOCATION: Nuclear.  
CC -I- TISSUE SPECIFICITY: EXPRESSED IN HEMATOPOIETIC AND NON  
CC HEMATOPOIETIC CELLS.  
CC -I- SIMILARITY: Contains 3 Myb-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X67505; CAA47839.1; -.  
CC PIR; S28050; S28050.  
CC PDB; 1A5U; 01-JUL-98.  
CC TRASNPAAC; T01586; -.  
CC InterPro; IPR001005; Myb DNA binding.  
CC Pfam; PF00249; myb DNA-binding; 3.  
DR SMART; SM00717; SANT; 3.  
DR PROSITE; PS00037; MYB 1; 3.  
DR PROSITE; PS00334; MYB 2; 3.  
DR PROSITE; PS50090; MYB 3; 3.  
KW Transcription regulation; Nuclear protein; DNA-binding; Repeat;  
3D-structure.  
FT DNA\_BIND 26 77 MYB 1.  
FT DNA\_BIND 78 129 MYB 2.  
FT DNA\_BIND 130 180 MYB 3.  
FT HELIX 88 101  
FT TURN 102 102  
FT HELIX 106 112  
FT HELIX 122 125  
FT TURN 126 126  
FT HELIX 140 147  
FT TURN 148 152  
FT HELIX 157 163  
FT TURN 165 166  
FT HELIX 169 178  
FT TURN 179 179  
FT TURN 183 184  
SQ SEQUENCE 686 AA; 77736 MW; 2D1209EAD648D7B CRC64;  
Query Match 17.3%; Score 255; DB 1; Length 686;  
Best Local Similarity 32.5%; Pred. No. 2.5e-11;  
Matches 63; Conservative 32; Mismatches 65; Indels 34; Gaps 5;  
QY 13 KGSWTDEDFLRKCIDKYGKWHLPVIRAGLNRCKRSLRLMYLRPHIKRGDFEOD 72  
DB 83 KGPWTKEDQKVIKLVKYGKWHLPVIRAGLNRCKRSLRLMYLRPHIKRGDFEOD 141  
QY 73 EVDILRLHLKLGKRWSLIAGRLPGRTANDVKNYNTNLRLKNTTKIVPREKINCKG 132  
DB 142 EDRITCEAHKVLGNRMALIAKMLPGRTDNAVKNHNSITKRVDTGGLNETK----- 194  
QY 133 ISTKIEIKPQRRKYSSTMKVNTNNVILDEEHCKEIISEKQTPASNDNDPMTINL 192  
DB 195 ESQPLYL-----VEVDNE-----SOSGTRAESQTIYVNMPEVDI 229  
QY 193 LENCNDDEED 206  
DB 230 SEIKEDV-SDEEV 242  
RESULT 15  
MYB\_DROME  
ID MYB\_DROME STANDARD; PRT; 657 AA.  
AC P04197; Q9VXK9;  
DT 20-MAR-1987 (Rel. 04, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myb protein.



GN MYB OR CG9045.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preyigota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 CC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Oregon-R;  
 RX MEDLINE=86082681; PubMed=3121304;  
 RA Peters C.W.B., Sippel A.E., Vingron M., Klemmner K.-H.;  
 RT "Drosophila and vertebrate myb proteins share two conserved regions,  
 RT one of which functions as a DNA-binding domain.";  
 RL EMBL J. 6:3085-3090(1987).  
 RP [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blaise J.R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Decher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dudbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,  
 RA Fobel C., Gargiela A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Goddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,  
 RA Jajani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lavoie P., Lei Y., Levytsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclio J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195 (2000).  
 RN [3]  
 RP SEQUENCE OF 1-441 FROM N.A.  
 RX MEDLINE=85176969; PubMed=3921261;  
 RA Katzen A.L., Kornberg T.B., Bishop J.M.;  
 RT "Isolation of the proto-oncogene c-myc from D. melanogaster.";  
 RL Cell 41:449-456(1985).  
 CC -1- FUNCTION: DNA-BINDING PROTEIN THAT SPECIFICALLY RECOGNIZES THE  
 CC -1- SEQUENCE 5'-TAAC(G/T)G-3'.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: Contains 3 Myb-like domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 CC EMBL: X05939; CA29373.1; -  
 CC EMBL: AE003500; AA048529.1; -  
 CC EMBL: M11281; AA070367.1; -  
 CC PIR: S00578; TRFMA.  
 CC HSSP: P01103; IPOM.  
 CC FLYBase: FBgn0002914; Myb.  
 CC GO: GO:0008283; P:cell proliferation; IMP.  
 CC GO: GO:0007098; P:centrosome cycle; IMP.  
 CC InterPro: IPR001005; Myb\_DNA\_binding.  
 CC Pfam: PF00249; myb\_DNA-binding; 3.  
 CC SMART: SM00717; SANT; 3.  
 CC PROSITE: PS00037; MYB\_1; 3.  
 CC PROSITE: PS00334; MYB\_2; 3.  
 CC PROSITE: PS50090; MYB\_3; 3.  
 CC Nucleic acid binding protein; DNA-binding; Repeat.  
 CC Nucleic acid binding protein; MYB 1.  
 CC Nucleic acid binding protein; MYB 2.  
 CC Nucleic acid binding protein; MYB 3.  
 CC CONFLICT 370 370 A -> V (IN REF. 1).  
 CC CONFLICT 440 441 QL -> KV (IN REF. 3).  
 CC SEQUENCE 657 AA; 74045 MW; 8265B37ABB250AE4 CRC64;  
 Query Match 17.1%; Score 252; DB 1; Length 657;  
 Best Local Similarity 33.5%; Pred. No. 3.9e-11;  
 Matches 57; Conservative 36; Mismatches 65; Indels 12; Gaps 4;  
 QY 13 KGSVTDEDFLLKRCIDKRGCKWHLVPIRAGLNCRKSKGRMLNLRPHIKRGDFED 72  
 DB 136 KGPWTRDEDDMVITLVNFGPKWTLI-ARYLNGRIQCRERHMHNLNPIKKTAWTEK 194  
 QY 73 EVDILRLHKLGRWELIAGRLPGRPTANDYKYNVNTNLKLTNTKIVPREKINKKGE 132  
 DB 195 EDLITVYAHLELGNQMAKIAKRLDGRDNDNAIKHNMSTMRKV-VERRSVNAAGSD 250  
 QY 133 I-STKIEIIRKRRKPSSTWKNVNTNNVILDEEHCKEIISEKQTPDAS 181  
 DB 251 LKSSRTHLITLTKSGGSKCKNNQNH-----KESGEAVNKENDGGA 294

Search completed: January 29, 2004, 20:11:54  
 Job time : 20 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 20:08:41 ; Search time 41 Seconds  
(without alignments)  
1724.547 Million cell updates/sec

Title: US-10-033-190-2  
Perfect score: 1477  
Sequence: 1 MNSTMSLIGVRKGSWTDEE.....QQGVNDPFSAEIDLWNLLD 274

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	694.5	47.0	255	10 Q9M721	Q9M721 petunia hyb
2	688.5	46.6	255	10 Q9M720	Q9M720 petunia hyb
3	682.5	46.2	255	10 Q9M719	Q9M719 petunia int
4	555	37.6	127	10 Q9M717	Q9M717 petunia hyb
5	551	37.3	129	10 Q9M718	Q9M718 petunia hyb
6	547	37.0	127	10 Q9M716	Q9M716 petunia hyb
7	543.5	36.8	247	10 Q8L5P2	Q8L5P2 vitis labru
8	542.5	36.7	246	10 Q9ENV9	Q9ENV9 arabidopsis
9	541	36.6	250	10 Q8L5P3	Q8L5P3 vitis labru
10	540	36.6	129	10 Q9M715	Q9M715 petunia axi
11	539.5	36.5	307	10 Q8L5P1	Q8L5P1 vitis labru
12	536	36.3	120	10 Q9M714	Q9M714 petunia axi
13	535	36.2	248	10 Q9P225	Q9P225 arabidopsis
14	532.5	36.1	249	10 Q9ZTC3	Q9ZTC3 arabidopsis
15	524	35.5	212	10 Q9ZTC7	Q9ZTC7 arabidopsis
16	503.5	34.1	139	10 Q9FNV8	Q9FNV8 arabidopsis

17	438.5	29.7	271	10 Q9ATD3	Q9ATD3 gossypium h
18	433	29.3	247	10 Q9ATD9	Q9ATD9 gossypium h
19	428	29.0	254	10 Q9ATD2	Q9ATD2 gossypium h
20	423.5	28.7	249	10 Q38850	Q38850 arabidopsis
21	415	28.1	302	10 Q9ATD5	Q9ATD5 gossypium h
22	409	27.7	272	10 Q9ZRS0	Q9ZRS0 oryza sativ
23	408	27.6	269	10 Q42379	Q42379 arabidopsis
24	405.5	27.5	388	10 Q40920	Q40920 picea maria
25	404	27.4	284	10 Q41869	Q41869 zea mays (m
26	403	27.3	218	10 Q8H253	Q8H253 gossypioide
27	403	27.3	226	10 Q8H255	Q8H255 gossypium r
28	403	27.3	246	10 Q9S7Y2	Q9S7Y2 arabidopsis
29	403	27.3	257	10 Q9S9K9	Q9S9K9 arabidopsis
30	402.5	27.3	226	10 Q49021	Q49021 gossypium h
31	402.5	27.3	226	10 Q8H256	Q8H256 gossypium h
32	401	27.1	203	10 Q9SE10	Q9SE10 arabidopsis
33	400	27.1	193	10 Q49018	Q49018 gossypium h
34	399.5	27.0	193	10 Q94FG3	Q94FG3 gossypium h
35	399.5	27.0	264	10 Q9XIU9	Q9XIU9 glycine max
36	397	26.9	266	10 Q41868	Q41868 zea mays (m
37	397	26.9	267	10 Q22450	Q22450 zea mays (m
38	397	26.9	271	10 Q41842	Q41842 zea mays (m
39	397	26.9	271	10 Q8GRW4	Q8GRW4 zea mays (m
40	397	26.9	342	10 Q9FJ07	Q9FJ07 arabidopsis
41	396	26.8	113	10 Q93V39	Q93V39 zea luxuria
42	396	26.8	181	10 Q8H262	Q8H262 gossypioide
43	396	26.8	198	10 Q49017	Q49017 gossypium h
44	396	26.8	254	10 Q9ATD1	Q9ATD1 gossypium h
45	395.5	26.8	184	10 Q94FG5	Q94FG5 gossypium r

## ALIGNMENTS

### RESULT 1

ID Q9M721 PRELIMINARY; PRT; 255 AA.  
AC Q9M721;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE An2 protein.  
GN AN2.  
OS Petunia hybrida (Petunia).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; lamiales; Solanales; Solanaceae; Petunia.  
OX NCBI\_TaxID=4102;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV\_V26; TISSUE=Petal limb;  
RX MEDLINE=9380006; PubMed=10449578;  
RA Quattrone F., Wing J., van der Woude K., Souer E., de Vetten N.,  
RA Mol J., Koes R.;  
RT "Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color.";  
RT Plant Cell 11:1433-1444(1999).  
RL CC - SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC - SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
CC EMBL; AFI46702; AAF66727.1; -  
DR HSSP; P06876; IMBK.  
DR TRANSFAC; T02955; -  
DR InterPro; IPR001005; MYB\_DNA\_binding.  
DR Pfam; PF00249; MYB\_DNA\_binding; 2.  
DR SMART; SM00717; SANT; 2.  
DR PROSITE; PS00037; MYB\_1; 1.  
DR PROSITE; PS00334; MYB\_2; 1.  
DR PROSITE; PS50090; MYB\_3; 2.  
KW DNA-binding; Nuclear Protein.  
SQ SSQUNCE 255 AA; 29021 MW; 3C590B8473209A52 CRC64;  
Query Match 47.0%; Score 694.5; DB 10; Length 255;  
Best Local Similarity 53.7%; Pred. No. 6.2e-47;

Matches 154; Conservative 33; Mismatches 53; Indels 47; Gaps 11;

QY 1 MNSTMSLSIGVRKSGWTEDEDFLLRKCTDKYGEKGMHLVPIRAGLNCRKSCRLRMWNTL 60  
Db 1 MSTNASTSGVRKAGWTEDEDFLLRECEIEKYGEKGMHLVPIRAGLNCRKSCRLRMWNTL 60

QY 61 RPHIKRGDFEODEVDLLRLHKLGNRWSLJAGRLPGRTANDVKNYNTNLRLKLTNTTKI 120  
Db 61 RPHIKRGDFSLDEVDLLRLHKLGNRWSLJAGRLPGRTANDVKNYNTNLRLKLTNTTKI 116

QY 121 VP---REKINNKCEISTKIEIIPQRRKFFSSTMKVNT---TNNVILDEEHCKE 170  
Db 117 APHDQKQESKKKAVKI-TENNIIKRPRTFSRPMANNFPCWNGKSCNKNTI-DKNEGDT 174

QY 171 II---SEKQTPDASMDVNDPWWINLNCNDIEDEEVVINYKTLTSLHBEISPLN 227  
Db 175 IIRKSDERKQKEESIDDLQWMANLNLN---NIEIEELVSCN---SPTLHBEETASVN 227

QY 228 IGEKNSMOQGOISHENWGEFSLNLPWMOQGVONDDFSAEI-DLWNL 273  
Db 228 --AESSLTQEG-----GGSLDSFSDIDIDLV 254

RESULT 2

ID Q9M720 PRELIMINARY; PRT; 255 AA.

AC Q9M720; 01-OCT-2000 (TRENBLrel. 15, Created)

DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE AN2 protein.

GN AN2.

OS Petunia integrifolia (Violet-flowered petunia) (Petunia inflata).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; lamids; Solanales; Solanaceae; Petunia.

OX NCBI\_TaxID=4103;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. S9; TISSUE=Petal limb;

RX MEDLINE=9938006; PubMed=10449578;

RA Quattrocchio F., Wing J., van der Woude K., Souer E., de Vetten N., Mol J., Koes R.;

RT "Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color.";

RL Plant Cell 11:1433-1444(1999).

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.

EMBL: AF146703; AAF66728.1; -.

DR HSSP: P06876; IMEX.

DR InterPro: IPR001005; MYB DNA binding.

DR Pfam: PF00249; MYB DNA-binding; 2.

DR SMART: SM00717; SANT; 2.

DR PROSITE: PS00037; MYB\_1; 1.

DR PROSITE: PS00334; MYB\_2; 1.

DR PROSITE: PS00090; MYB\_3; 2.

KW DNA-binding; Nuclear protein.

SEQUENCE 255 AA; 29007 MW; 48818508AFD59C3F CRC64;

Query Match 46.6%; Score 688.5; DB 10; Length 255;  
Best Local Similarity 53.0%; Pred. No. 1.8e-46;  
Matches 152; Conservative 37; Mismatches 51; Indels 47; Gaps 12;

QY 1 MNSTMSLSIGVRKSGWTEDEDFLLRKCTDKYGEKGMHLVPIRAGLNCRKSCRLRMWNTL 60  
Db 1 MSTNASTSGVRKAGWTEDEDFLLRECEIEKYGEKGMHLVPIRAGLNCRKSCRLRMWNTL 60

QY 61 RPHIKRGDFEODEVDLLRLHKLGNRWSLJAGRLPGRTANDVKNYNTNLRLKLTNTTKI 120  
Db 61 RPHIKRGDFSLDEVDLLRLHKLGNRWSLJAGRLPGRTANDVKNYNTNLRLKLTNTTKI 116

QY 121 VP---REKINNKCEISTKIEIIPQRRKFFSSTMKVNT---TNNVILDEEHCKE 170  
Db 117 APHDQKQESKKKAVKI-TENNIIKRPRTFSRPMANNFPCWNGKSCNKNTI-DKNEGDT 174

QY 171 II---SEKQTPDASMDVNDPWWINLNCNDIEDEEVVINYKTLTSLHBEISPLN 227  
Db 175 IIRKSDERKQKEESIDDLQWMANLNLN---NIEIEELVSCN---SPTLHBEETASVN 227

QY 228 IGEKNSMOQGOISHENWGEFSLNLPWMOQGVONDDFSAEI-DLWNL 273  
Db 228 --AESSLTQEG-----GGSLDSFSDIDIDLV 254

Db 117 APHDQKQESKKKAVKI-TENSIIPKRPRTFSRPMANNFPCWNGKSCNKNTI-DKNEGDT 174

QY 171 II---SEKQTPDASMDVNDPWWINLNCNDIEDEEVVINYKTLTSLHBEISPLN 227  
Db 175 IIRKSDERKQKEESIDDLQWMANLNLN---NIEIEELVSCN---SPTLHBEETASVN 227

QY 228 IGEKNSMOQGOISHENWGEFSLNLPWMOQGVONDDFSAEI-DLWNL 273  
Db 228 --AESSLTQEG-----GGSLDSFSDIDIDLV 254

RESULT 3

ID Q9M719 PRELIMINARY; PRT; 255 AA.

AC Q9M719; 01-OCT-2000 (TRENBLrel. 15, Created)

DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE AN2 protein.

GN AN2.

OS Petunia integrifolia (Violet-flowered petunia) (Petunia inflata).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; lamids; Solanales; Solanaceae; Petunia.

OX NCBI\_TaxID=4103;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. S6; TISSUE=Petal limb;

RX MEDLINE=9938006; PubMed=10449578;

RA Quattrocchio F., Wing J., van der Woude K., Souer E., de Vetten N., Mol J., Koes R.;

RT "Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color.";

RL Plant Cell 11:1433-1444(1999).

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.

EMBL: AF146704; AAF66729.1; -.

DR HSSP: P06876; IMEX.

DR InterPro: IPR001005; MYB DNA binding.

DR Pfam: PF00249; MYB DNA-binding; 2.

DR SMART: SM00717; SANT; 2.

DR PROSITE: PS00037; MYB\_1; 1.

DR PROSITE: PS00090; MYB\_3; 2.

KW DNA-binding; Nuclear protein.

SEQUENCE 255 AA; 28963 MW; E4626F08B07071AD CRC64;

Query Match 46.2%; Score 682.5; DB 10; Length 255;  
Best Local Similarity 52.3%; Pred. No. 5.5e-46;  
Matches 150; Conservative 39; Mismatches 51; Indels 47; Gaps 11;

QY 1 MNSTMSLSIGVRKSGWTEDEDFLLRKCTDKYGEKGMHLVPIRAGLNCRKSCRLRMWNTL 60  
Db 1 MSTNASTSGVRKAGWTEDEDFLLRECEIEKYGEKGMHLVPIRAGLNCRKSCRLRMWNTL 60

QY 61 RPHIKRGDFEODEVDLLRLHKLGNRWSLJAGRLPGRTANDVKNYNTNLRLKLTNTTKI 120  
Db 61 RPHIKRGDFSLDEVDLLRLHKLGNRWSLJAGRLPGRTANDVKNYNTNLRLKLTNTTKI 116

QY 121 VP---REKINNKCEISTKIEIIPQRRKFFSSTMKVNT---TNNVILDEEHCKE 170  
Db 117 APHDQKQESKKKAVKI-TENNIIKRPRTFSRPMANNFPCWNGKSCNKNTI-DKNEGDT 174

QY 171 II---SEKQTPDASMDVNDPWWINLNCNDIEDEEVVINYKTLTSLHBEISPLN 227  
Db 175 IIRKSDERKQKEESIDDLQWMANLNLN---NIEIEELVSCN---SPTLHBEETASVN 227

QY 228 IGEKNSMOQGOISHENWGEFSLNLPWMOQGVONDDFSAEI-DLWNL 273  
Db 228 --AESSLTQEG-----GGSLDSFSDIDIDLV 254

RESULT 4

ID Q9M717

ID Q9M717 PRELIMINARY; PRT; 127 AA.  
 AC Q9M717;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE An2 truncated protein.  
 GN AN2.  
 OS Petunia hybrida (Petunia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamids; Solanales; Solanaceae; Petunia.  
 OC NCBI\_TaxID=4102;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. W22; TISSUE=Petal limb;  
 RX MEDLINE=99380006; PubMed=10449578;  
 RA Quattrocchio F., Wing J., van der Woude K., Souer E., de Vetten N.,  
 RA Mol J., Koes R.;  
 RT "Molecular analysis of the anthocyanin2 gene of petunia and its role  
 in the evolution of flower color.";  
 RL Plant Cell 11:1433-1444(1999).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
 DR EMBL; AFI46706; AAF66731.1; -.  
 DR HSSP; P06876; IMBK.  
 DR InterPro; IPR001005; MYB DNA binding.  
 DR Pfam; PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KW DNA-binding; Nuclear protein.  
 SQ SEQUENCE 127 AA; 14789 MW; 583D573C944A8E8 CRC64;

Query Match 37.6%; Score 555; DB 10; Length 127;  
 Best Local Similarity 80.0%; Pred. No. 2.6e-36;  
 Matches 100; Conservative 13; Mismatches 8; Indels 4; Gaps 1;  
 QY 1 MNSTSSSLGVRKGSWTEDEDFLLRKIDKYGSGKWLVPVRAGLNRCRSCRLRWNTLY 60  
 DB 1 MSTSSNSTSGVRKGATTEEDDLIRKIDKYGSGKWLVPVRAGLNRCRSCRLRWNTLY 60  
 QY 61 RPHIKRGDEQDEVDLILRLHKLGNRWSLIAGRLPGRANDVKNYNTNLRLKNTTKI 120  
 DB 61 RPHIKRGDEQDEVDLILRLHKLGNRWSLIAGRLPGRANDVKNYNTNLRLKNTTKI 120  
 QY 121 VPPEK 125  
 DB 117 APHDQ 121  
 RESULT 5  
 Q9M718 PRELIMINARY; PRT; 129 AA.  
 ID Q9M718;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE An2 truncated protein.  
 GN AN2.  
 OS Petunia hybrida (Petunia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamids; Solanales; Solanaceae; Petunia.  
 OC NCBI\_TaxID=4102;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. W115; TISSUE=Petal limb;  
 RX MEDLINE=99380006; PubMed=10449578;  
 RA Quattrocchio F., Wing J., van der Woude K., Souer E., de Vetten N.,  
 RA Mol J., Koes R.;  
 RT "Molecular analysis of the anthocyanin2 gene of petunia and its role  
 in the evolution of flower color.";

RL Plant Cell 11:1433-1444(1999).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
 DR EMBL; AFI46705; AAF66730.1; -.  
 DR HSSP; P06876; IMBK.  
 DR InterPro; IPR001005; MYB DNA binding.  
 DR Pfam; PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KW DNA-binding; Nuclear protein.  
 SQ SEQUENCE 129 AA; 15084 MW; 12D9D38441D2427 CRC64;  
 Query Match 37.3%; Score 551; DB 10; Length 129;  
 Best Local Similarity 79.2%; Pred. No. 5.4e-36;  
 Matches 99; Conservative 15; Mismatches 7; Indels 4; Gaps 1;  
 QY 1 MNSTSSSLGVRKGSWTEDEDFLLRKIDKYGSGKWLVPVRAGLNRCRSCRLRWNTLY 60  
 DB 1 MSTSSNSTSGVRKGATTEEDDLIRKIDKYGSGKWLVPVRAGLNRCRSCRLRWNTLY 60  
 QY 61 RPHIKRGDEQDEVDLILRLHKLGNRWSLIAGRLPGRANDVKNYNTNLRLKNTTKI 120  
 DB 61 RPHIKRGDEQDEVDLILRLHKLGNRWSLIAGRLPGRANDVKNYNTNLRLKNTTKI 120  
 QY 121 VPPEK 125  
 DB 117 VPHDQ 121

RESULT 6  
 Q9M716 PRELIMINARY; PRT; 127 AA.  
 ID Q9M716;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE An2 truncated protein.  
 GN AN2.  
 OS Petunia hybrida (Petunia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamids; Solanales; Solanaceae; Petunia.  
 OC NCBI\_TaxID=4102;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. W44; TISSUE=Petal limb;  
 RX MEDLINE=99380006; PubMed=10449578;  
 RA Quattrocchio F., Wing J., van der Woude K., Souer E., de Vetten N.,  
 RA Mol J., Koes R.;  
 RT "Molecular analysis of the anthocyanin2 gene of petunia and its role  
 in the evolution of flower color.";  
 RL Plant Cell 11:1433-1444(1999).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
 DR EMBL; AFI46707; AAF66732.1; -.  
 DR HSSP; P06876; IMBK.  
 DR InterPro; IPR001005; MYB DNA binding.  
 DR Pfam; PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KW DNA-binding; Nuclear protein.  
 SQ SEQUENCE 127 AA; 14817 MW; F928569624274F35 CRC64;  
 Query Match 37.0%; Score 547; DB 10; Length 127;  
 Best Local Similarity 78.4%; Pred. No. 1.1e-35;  
 Matches 98; Conservative 15; Mismatches 8; Indels 4; Gaps 1;  
 QY 1 MNSTSSSLGVRKGSWTEDEDFLLRKIDKYGSGKWLVPVRAGLNRCRSCRLRWNTLY 60  
 DB 1 MSTSSNSTSGVRKGATTEEDDLIRKIDKYGSGKWLVPVRAGLNRCRSCRLRWNTLY 60

```

Db      1 MSTSNASTGVRKAMTEEDILLRECIKYEKGKWLVPVAGLANCRKSCRLRWNL 60
Qy      61 RPHIKRDEPDEVDILRLHKLGNRWSLIAGLPRTANDVKNYNTNLRLKNTKI 120
Db      61 RPHIKRDEPDEVDILRLHKLGNRWSLIAGLPRTANDVKNYNTNLRLKNTKI 116
Qy      121 VPREK 125
Db      117 APHDQ 121

RESULT 7
Q8LSP2  PRELIMINARY; PRT; 247 AA.
AC      08LSP2;
DT      01-OCT-2002 (TREMBlrel. 22, Created)
DT      01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      MYB-related transcription factor VIMYBA1-2.
GN      VIMYBA1-2.
OS      Vitis labrusca x Vitis vinifera.
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC      Vitis.
OX      NCBI_TaxID=105599;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=cv. Kyoho; TISSUE=fruit;
RA      Kobayashi S., Ishimaru M., Hiraoka K., Honda C.;
RT      "A myb-related gene of the Kyoho grape (Vitis labruscana) regulates
RT      anthocyanin biosynthesis in grapes."
RL      submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC      -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC      -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
DR      EMBL; AB073012; BAC07539.1; -.
DR      InterPro; IPR001005; MYB DNA binding.
DR      Pfam; PF00249; myb DNA-binding; 2.
DR      SMART; SM00717; SANT; 2.
DR      PROSITE; PS50090; MYB_3; 2.
KW      DNA-binding; Nuclear protein.
SQ      SEQUENCE 247 AA; 28255 MW; 9AA1717840F8A8BD CRC64;

Query Match      36.8%; Score 543.5; DB 10; Length 247;
Best Local Similarity 45.3%; Pred. No. 4.6e-35;
Matches 121; Conservative 35; Mismatches 80; Indels 31; Gaps 8;

Qy      6 MSSIGVRKGSWTDEDFLLRKIDKYGEGKWLVPVIRAGLNCRKSCRLRWNLRLPHIK 65
Db      1 MESIGVRKGSWTDEDFLLRKIDKYGEGKWLVPVIRAGLNCRKSCRLRWNLRLPHIK 60
Qy      66 RGDPEQDEVDILRLHKLGNRWSLIAGLPRTANDVKNYNTNLRLKNTTKIVPREK 125
Db      61 RGEFALDEVDILRLHKLGNRWSLIAGLPRTANDVKNYNTNLRLKNTTKIVPREK 116
Qy      126 INNKGELSTKLEIKRQRYKFSSTKNTNNVNLDEEHEKELISEKOTPDASMDNV 185
Db      117 GRNK-PLTHSTKAIKPHPK-FSKALPRFELKTTAVDTP-TQVSTSSKSSSPQND 173
Qy      186 D-PWMINLENCNDIDEEDEEVINTEKTLTSLHEEISPLNTGEGNSMOGOISHEMW 244
Db      174 DIIMESILA-----ELDQETDFSASGEMLIASLRAEFAT-----QKKGPMDGMIETOG 224
Qy      245 GEFSLNLPMPQGVQNDPFAIDIMN 271
Db      225 GE-----GDFFPDVGFMD 237

RESULT 8
Q9FNV9  PRELIMINARY; PRT; 246 AA.
AC      Q9FNV9;
DT      01-MAR-2001 (TREMBlrel. 16, Created)
DT      01-MAR-2001 (TREMBlrel. 16, Last sequence update)

```

```

DT      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      Putative transcription factor MYB113 (Myb-related transcription
DE      factor, putative).
GN      T27F4.12.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      Stracke R., Weisshaar B.;
RT      "MYB transcription factor gene nomenclature in Arabidopsis thaliana."
RL      Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=cv. Columbia;
RX      MEDLINE=21016719; PubMed=11130712;
RA      Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA      White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA      Buehler E., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA      Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA      Dunn P., Elgu P., Feldlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA      Gall J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA      Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin B.,
RA      Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA      Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA      Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzali A.,
RA      Millefischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA      Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA      Sakano H., Salzer J.S., Schwartz J.R., Shinn P., Southwick A.M.,
RA      Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA      Utechtack T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA      Wu D., Yu G., Praser C.M., Venter J.C., Davis R.M.;
RT      "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT      thaliana."
RL      Nature 408:816-820(2000).
CC      -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC      -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
DR      EMBL; AY008378; AAC8380.1; -.
DR      EMBL; AC020665; AAC52158.1; -.
DR      HSP; P06876; IMB.
DR      InterPro; IPR001005; MYB DNA binding.
DR      Pfam; PF00249; myb DNA-binding; 2.
DR      SMART; SM00717; SANT; 2.
DR      PROSITE; PS00037; MYB_1; 1.
DR      PROSITE; PS00334; MYB_2; 1.
DR      PROSITE; PS50090; MYB_3; 2.
KW      DNA-binding; Nuclear protein.
SQ      SEQUENCE 246 AA; 28307 MW; FF1C4B9176040D20 CRC64;

Query Match      36.7%; Score 542.5; DB 10; Length 246;
Best Local Similarity 52.4%; Pred. No. 5.5e-35;
Matches 109; Conservative 26; Mismatches 52; Indels 21; Gaps 4;

Qy      7 SSGVRKGSWTDEDFLLRKIDKYGEGKWLVPVIRAGLNCRKSCRLRWNLRLPHIKR 66
Db      4 SPKGRKGTWTTEDDILRLHKLGNRWSLIAGLPRTANDVKNYNTNLRLKNTTKIVPREK 63
Qy      67 GDFEODEVDILRLHKLGNRWSLIAGLPRTANDVKNYNTNLRLKNTTKIVPREK 122
Db      64 GKLCSDVDVLRLHKLGNRWSLIAGLPRTANDVKNYNTNLRLKNTTKIVPREK 122
Qy      123 REKLNKCGELSTKLEIKRQRYKFSSTKNTNNVNLDEEHEKELISEKOTPDASMDNV 182
Db      123 NKNTSHPTSSAOKIDVLKPRPSFSDKNSCNVNIIPKVVVPLHLGLNNNYVCCSIT 182
Qy      168 C-KELISEKOTPDASMDNVPPWNLLE 194
Db      183 CNDEQKDKLINTNLGDGNNWMSLE 210

RESULT 9

```

```

Q8L5P3
ID 08L5P3 PRELIMINARY; PRT; 250 AA.
AC 08L5P3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Myb-related transcription factor VIMYBA1-1.
GN VIMYBA1-1.
OS Vitis labrusca x Vitis vinifera.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC Vitis.
OX NCBI_TaxID=105599;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Kyoho; TISSUE=fruit;
RA Kobayashi S., Ishimaru M., Hiraoka K., Honda C.;
RT "A myb-related gene of the Kyoho grape (Vitis labruscana) regulates
RT anthocyanin biosynthesis in grapes.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
DR EMBL: AB073010; BAC07537.1; -.
DR InterPro: IPR001005; Myb DNA binding.
DR Pfam: PF00249; myb DNA-binding; 2.
DR SMART: SM00717; SANT; 2.
DR PROSITE: PS50090; MYB_3; 2.
DR DNA-binding: Nuclear Protein.
KM SEQUENCE 250 AA; 28597 MW; 9DA7740CE0A99DEA CRC64;
SQ
Query Match 36.6%; Score 541; DB 10; Length 250;
Best Local Similarity 44.9%; Pred. No. 7,4e-35;
Matches 120; Conservative 37; Mismatches 82; Indels 28; Gaps 8;

QY 6 MSSLGVRKGSWTDDEDFLLRKCDIKYGEKWHLPVIRAGLNRCRKSRLRWMLNLYLRPHIK 65
DB 1 MSSLGVRKGMWIOEDVLLRKCIKYEKGKWHLPVIRAGLNRCRKSRLRWMLNLYLRPHIK 60
QY 66 RSGPDEDEVLLIRHLKLNRMWSLNGRLPGRTANVKNYWNNTLRLKNTTKIVPREK 125
DB 61 RSGPDEDEVLLIRHLKLNRMWSLNGRLPGRTANVKNYWNNTLRLKNTTKIVPREK 116
QY 126 INKCGISTKIEIKPQRKRYFSSITMKNTNNVILDEBEBKEIISEKOTPDASMDNV 185
DB 117 GRKK-PQTHSKTKAIKHPK-PKALPKPEIKTAVDTDP-QVSTSSRPSSTSPEND 173
QY 186 D-PWMLNLNLCNDDEIEDEEVVINYKTLTSLHBEISPLNIGEGNSMQGQISHEW 244
DB 174 DIIMESILAE-HAQMQDETDFASGEMLIASLAESTAT-----QKGPMDGWIETQIG 227
QY 245 GEFSLNLPWQGVQNDPSAIDLMN 271
DB 228 GE-----GDFPFDGVGFWMD 240

RESULT 10
Q9M715 PRELIMINARY; PRT; 129 AA.
AC Q9M715;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE An2 truncated protein.
GN AN2.
OS Petunia axillaris.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanales; Solanaceae; Petunia.
OX NCBI_TaxID=33119;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. SL; TISSUE=petal limb;
RX MEDLINE=99380006; Pubmed=10449578;

```

```

RA Quattrocchio F., Wang J., van der Woude K., Souer E., de Vetten N.,
RA Mol J., Koes R.;
RT "Molecular analysis of the anthocyanin2 gene of petunia and its role
RT in the evolution of flower color.";
RL Plant Cell 11:1433-1444(1999).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
DR EMBL: AF146708; AAF66733.1; -.
DR HSSP: P06876; 1MBK.
DR InterPro: IPR001005; Myb DNA binding.
DR Pfam: PF00249; myb DNA-binding; 2.
DR SMART: SM00717; SANT; 2.
DR PROSITE: PS00037; MYB_1; 1.
DR PROSITE: PS00334; MYB_2; 1.
DR PROSITE: PS50090; MYB_3; 2.
KM DNA-binding: Nuclear Protein.
SQ SEQUENCE 129 AA; 15041 MW; 0E3731FDD84FDIC7 CRC64;

Query Match 36.6%; Score 540; DB 10; Length 129;
Best Local Similarity 76.8%; Pred. No. 4e-35;
Matches 96; Conservative 18; Mismatches 7; Indels 4; Gaps 1;

QY 1 MNSTMSGLGVRKGSWTDDEDFLLRKCDIKYGEKWHLPVIRAGLNRCRKSRLRWMLNLY 60
DB 1 MSTMSNSTSGVRKGMWTEEDLILRRCIEKYGEKWHLPVIRAGLNRCRKSRLRWMLNLY 60
QY 61 RPHIKRGDFEDEVLLIRHLKLNRMWSLNGRLPGRTANVKNYWNNTLRLKNTTKI 120
DB 61 RPHIKRGDFEDEVLLIRHLKLNRMWSLNGRLPGRTANVKNYWNNTLRLKNTTKI 116
QY 121 VPREK 125
DB 117 VPHDQ 121

RESULT 11
Q8L5P1 PRELIMINARY; PRT; 307 AA.
AC 08L5P1;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Myb-related transcription factor VIMYBA2.
GN VIMYBA2.
OS Vitis labrusca x Vitis vinifera.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC Vitis.
OX NCBI_TaxID=105599;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Kyoho; TISSUE=fruit;
RA Kobayashi S., Ishimaru M., Hiraoka K., Honda C.;
RT "A myb-related gene of the Kyoho grape (Vitis labruscana) regulates
RT anthocyanin biosynthesis in grapes.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
DR EMBL: AB073013; BAC07540.1; -.
DR InterPro: IPR001005; Myb DNA binding.
DR Pfam: PF00249; myb DNA-binding; 2.
DR SMART: SM00717; SANT; 2.
DR PROSITE: PS50090; MYB_3; 2.
DR DNA-binding: Nuclear Protein.
KM SEQUENCE 307 AA; 34891 MW; 682ACD358C3007BE CRC64;

Query Match 36.5%; Score 539.5; DB 10; Length 307;
Best Local Similarity 46.6%; Pred. No. 1.2e-34;
Matches 124; Conservative 35; Mismatches 88; Indels 19; Gaps 8;

QY 6 MSSLGVRKGSWTDDEDFLLRKCDIKYGEKWHLPVIRAGLNRCRKSRLRWMLNLYLRPHIK 65
DB 1 MSSLGVRKGMWIOEDVLLRKCIKYEKGKWHLPVIRAGLNRCRKSRLRWMLNLYLRPHIK 60

```

```

QY 66 RGDFEODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLRLKNTTKIVPREK 125
DB 61 RGEALDEVDLMTLNLGNRWSLIAGRLPGRTANDVKNYNTNLRLKNTTKIVPREK 116
QY 126 INNKGESTIKETIKORRKYFSGTMKNVNNVILDEEHCHEIISEKOTPDASMDNV 185
DB 117 GRDK-PQTHSKTKAIKPHPK-FSKALPKFELKTTAVDTFD-TQVSTSSKPSSTSPQND 173
QY 186 D-PWIMILNENCDIDEEDEAVINYEKTLTSLHBEISPLINIGENSMQOQISHENW 244
DB 174 DIIMWESLAE-HAOMQDETDFSSAGELILASLAEETAT-----QKKGPMGMEIOIG 227
QY 245 GE-----FSLNLPMPQGVQNDPFA 265
DB 228 GEGDIWESLAEHAQMDQDTPFA 253

RESULT 12
ID 09M714 PRELIMINARY; PRT; 120 AA.
AC 09M714;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE An2 truncated protein.
CN AN2.
OS Petunia axillaris.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanales; Solanaceae; Petunia.
OC NCBI_TaxID=33119;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. S7; TISSUE=Petal limb;
RX MEDLINE=9938006; PubMed=10449578;
RA Quattrocchio F., Wing J., van der Woude K., Souer E., de Vetten N.,
RA Mol J., Koes R.;
RT "Molecular analysis of the anthocyanin2 gene of petunia and its role
RT in the evolution of flower color.";
RL Plant Cell 11:1433-1444(1999).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
EMBL: AF146709; AAF6734.1; -.
HSSP: P06876; IMK.
DR InterPro: IPR001005; Myb_DNA_binding.
DR Pfam: PF00249; myb_DNA-binding; 2.
DR SMART: SM00717; SANT; 2.
DR PROSITE: PS00037; MYB_1; 1.
DR PROSITE: PS00334; MYB_2; 1.
DR PROSITE: PS50090; MYB_3; 2.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 120 AA; 13935 MW; AEEED2648C9726EFC CRC64;

Query Match 36.3%; Score 536; DB 10; Length 120;
Best Local Similarity 79.0%; Pred. No. 7.5e-35;
Matches 98; Conservative 13; Mismatches 9; Indels 4; Gaps 1;

```

```

ID 09FE25 PRELIMINARY; PRT; 248 AA.
AC 09FE25;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Production of anthocyanin pigment 1 protein (putative transcription
DE factor).
CN PAPI OR P25P12.92.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosidae;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia-0;
RA Borevitz J.O., Xia Y., Blount J., Dixon R.A., Lamb C.;
RT "Activation Tagging Identifies a Conserved MYB Regulator of
RT Phenylpropanoid Biosynthesis.";
RL Plant Cell 12:0-0(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Federpiehl N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Nguyen M., Lam B., Southwick A., Bei, Buehler E., Chin C.,
RA Chou J., Choi E., Dunn P., Gonzalez A., Huang B., Kim C., Koo T.,
RA Lee J.M., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
RA Pham P., Sakano H., Schwartz J., Shinn P., Thaveri A., Toriumi M.,
RA Vaysberg M., Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
EMBL: AF325123; AAG42001.1; -.
DR EMBL: AC009323; AAG09100.1; -.
DR HSSP: P06876; IIDY.
DR InterPro: IPR001005; Myb_DNA_binding.
DR Pfam: PF00249; myb_DNA-binding; 2.
DR SMART: SM00717; SANT; 2.
DR PROSITE: PS00037; MYB_1; 1.
DR PROSITE: PS00334; MYB_2; 1.
DR PROSITE: PS50090; MYB_3; 2.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 248 AA; 28469 MW; B86103A9ABSCF442 CRC64;

Query Match 36.2%; Score 535; DB 10; Length 248;
Best Local Similarity 44.2%; Pred. No. 2.2e-34;
Matches 122; Conservative 40; Mismatches 68; Indels 46; Gaps 10;

```

DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)  
 DE 01-MAR-2003 (TREMBlrel. 23, last annotation update)  
 DE Putative transcription factor (Production of anthocyanin pigment 2  
 DE protein) (MYB-related transcription factor, putative, 65699-  
 DE 67047).  
 GN MYB90 OR PAP2 OR T27F4.14.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=99056848; PubMed=9839469;  
 RA Kranz H.D., Denkamp M., Greco R., Jin H.-L., Leyva A., Meissner R.,  
 RA Petroni K., Urzainqui A., Beyer M., Martin C., Smeekens S.,  
 RA Tonelli C., Paz-Ares J., Weissshaar B.;  
 RT "Towards functional characterisation of the members of the R2R3-MYB  
 RT gene family from Arabidopsis thaliana.";  
 RL Plant J. 16:263-276(1998).  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia-0;  
 RA Borevitz J.O., Xia Y., Blount J., Dixon R.A., Lamb C.;  
 RT "Activation Tagging identifies a Conserved MYB Regulator of  
 RT Phenylpropanoid Biosynthesis.";  
 RL Plant Cell 12:0-0(2000).  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Benlier E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gail J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kuritz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,  
 RA Miltchev J., Miranda M., Nguyen M., Nierman W.C., Osborne B.T.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utecher T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 408:816-820(2000).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
 DR EMBL: AF062915; AAC83637.1; -;  
 DR EMBL: AF325124; AAC42002.1; -;  
 DR EMBL: AC020665; AAC52164.1; -;  
 DR HSSP: P06876; 1IDY.  
 DR InterPro: IPR001005; MYB\_DNA\_binding.  
 DR Pfam: PF00249; myb\_DNA-binding; 2.  
 DR SMART: SM00717; SANT; 2.  
 DR PROSITE: PS00037; MYB\_1; 1.  
 DR PROSITE: PS00334; MYB\_2; 1.  
 DR PROSITE: PS50090; MYB\_3; 2.  
 KW DNA-binding; Nuclear protein.  
 SQ SEQUENCE 249 AA; 28100 MW; 309024F04103700A CRC64;  
 Query March 36.1%; Score 532.5; DB 10; Length 249;  
 Best Local Similarity 44.4%; Pred. No. 3.4e-34;  
 Matches 122; Conservative 34; Mismatches 76; Indels 43; Gaps 7;  
 QY 7 SSGAGKGSWTDDEDFLLRKICIDYKGGKWHLPVRAGLNRCKSCRLRWLNTYLRPIK 66  
 DB 4 SSKGLRGKAWTAEEDSLRLCLIDKYGKQHWQVPLRAGLNRCKSCRLRWLNTYLRPIK 63

QY 67 GDEODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLARKLNTYTKIVPREKI 126  
 DB 64 GRLSNDSVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLARKLNTYTKIVPREKI 123  
 QY 127 NNKGEIST--KIEIKPQRKRYFSSMTKQNTNNVILDEBEHCKEIISEKQTPDASMD 183  
 DB 124 KNISPTTTPVQKIGVFKPRPSF---SVNNGCSHLNGLPEVDLIPSLGLKKN----- 174  
 QY 184 NVDPWNTNLEN---CNDDIEDEEVVINYKTLTSLHBEISPPNLIGSGNSMQGQIS 240  
 DB 175 -----NVCENSITCKDKDEKDFV-----NNLWNGDNWMLLENL 208  
 QY 241 HENWGEFSLKLPMPQGVQNDPFAEID-LNLTLD 274  
 DB 209 GEN-QEADALVPERTTAEHATITLAFVDEQJMSLFD 242  
 RESULT 15  
 ID 09ZTC7 PRELIMINARY; PRT; 212 AA.  
 AC 09ZTC7;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)  
 DE 01-MAR-2003 (TREMBlrel. 23, last annotation update)  
 DE Putative transcription factor.  
 GN MYB75  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=99056848; PubMed=9839469;  
 RA Kranz H.D., Denkamp M., Greco R., Jin H., Leyva A., Meissner R.C.,  
 RA Petroni K., Urzainqui A., Beyer M., Martin C., Smeekens S.,  
 RA Tonelli C., Paz-Ares J., Weissshaar B.;  
 RT "Towards functional characterisation of the members of the R2R3-MYB  
 RT gene family from Arabidopsis thaliana.";  
 RL Plant J. 16:263-276(1998).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
 DR EMBL: AF062908; AAC83630.1; -;  
 DR HSSP: P06876; 1IDY.  
 DR InterPro: IPR001005; MYB\_DNA\_binding.  
 DR Pfam: PF00249; myb\_DNA-binding; 2.  
 DR SMART: SM00717; SANT; 2.  
 DR PROSITE: PS00037; MYB\_1; 1.  
 DR PROSITE: PS00334; MYB\_2; 1.  
 DR PROSITE: PS50090; MYB\_3; 2.  
 KW DNA-binding; Nuclear protein.  
 SQ SEQUENCE 212 AA; 24508 MW; DE53594D17A0A240 CRC64;  
 Query March 35.5%; Score 524; DB 10; Length 212;  
 Best Local Similarity 48.3%; Pred. No. 1.3e-33;  
 Matches 102; Conservative 33; Mismatches 38; Indels 38; Gaps 5;  
 QY 7 SSGAGKGSWTDDEDFLLRKICIDYKGGKWHLPVRAGLNRCKSCRLRWLNTYLRPIK 66  
 DB 4 SSKGLRGKAWTAEEDSLRLCLIDKYGKQHWQVPLRAGLNRCKSCRLRWLNTYLRPIK 63  
 QY 67 GDEODEVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLARKLNTYTKIVPREKI 126  
 DB 64 GRLSNDSVDLILRLHKLGNRWSLIAGRLPGRTANDVKNYNTNLARKLNTYTKIVPREKI 123  
 QY 127 NNKGEIST--KIEIKPQRKRYFSSMTKQNTNNVILDEBEHCKEIISEKQTPDASMD 183  
 DB 124 KNISPTTTPVQKIGVFKPRPSF---SVNNGCSHLNGLPEVDLIPSLGLKKN----- 174  
 QY 184 NVDPWNTNLEN---CNDDIEDEEVVINYKTLTSLHBEISPPNLIGSGNSMQGQIS 240  
 DB 175 -----NVCENSITCKDKDEKDFV-----NNLWNGDNWMLLENL 208  
 QY 241 HENWGEFSLKLPMPQGVQNDPFAEID-LNLTLD 274  
 DB 209 GEN-QEADALVPERTTAEHATITLAFVDEQJMSLFD 242  
 QY 160 VILDEEHCKEIISEKQTPDASMDNVDPMWI 190  
 DB 124 DTFPIPTPALKNVYKPRPSFTVANNDCNHLNAPKVDVNP---CLGLNTNNVCDNS 179

Mon Feb 2 12:05:37 2004

us-10-033-190-2.rspt

Page 8

Db 180 IYNNKKKKQOLVN-----NLIDGDNMML 203

Search completed: January 29, 2004, 20:12:48  
Job time : 43 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 20:13:27 ; Search time 41 Seconds  
(without alignments)  
1060.758 Million cell updates/sec

Title: US-10-033-190-2

Sequence: 1 NMSTSMSSLGVRKSWTDEE.....QQGVNDPFSAEIDLWNLID 274

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

A\_Geneseq\_19Jun03:\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	274	100.0	274	23	AB811626
2	32	11.7	209	22	AA801933
3	32	11.7	209	23	AAU93176
4	32	11.7	246	22	AA801932
5	32	11.7	246	23	AAU93177
6	32	11.7	249	22	AA801916
7	32	11.7	249	22	AA868356
8	32	11.7	249	23	AAU92969
9	32	11.7	249	23	AAU75735

10	32	11.7	249	24	AA830046	Arabidopsis thalia
11	27	9.9	211	22	AA868355	Amino acid sequenc
12	27	9.9	212	22	AAU93155	Arabidopsis transc
13	27	9.9	248	23	AAU75734	A. thaliana Produc
14	20	7.3	302	22	AA882483	Cotton transcripti
15	17	6.2	48	21	AA833327	Eucalyptus grandis
16	17	6.2	65	21	AA833337	Eucalyptus grandis
17	17	6.2	109	21	AA833386	Eucalyptus grandis
18	17	6.2	113	21	AA833050	Pinus radiata tran
19	17	6.2	120	23	ABJ10409	Myb-related transc
20	17	6.2	128	21	AA833034	Pinus radiata tran
21	17	6.2	131	21	AA818501	Pinus radiata tran
22	17	6.2	145	21	AA818501	Pinus radiata tran
23	17	6.2	153	21	AA833239	Arabidopsis thalia
24	17	6.2	181	21	AA816627	Arabidopsis thalia
25	17	6.2	195	21	AA816626	Arabidopsis thalia
26	17	6.2	226	22	AA882482	Cotton transcripti
27	14	5.1	68	21	AA833401	Pinus radiata tran
28	14	5.1	89	21	AA833256	Pinus radiata tran
29	14	5.1	104	21	AA833028	Pinus radiata tran
30	14	5.1	203	23	AAU93131	Arabidopsis transc
31	14	5.1	219	23	AAU92997	Arabidopsis transc
32	14	5.1	226	21	AA833301	Pinus radiata tran
33	14	5.1	231	22	AA882484	Cotton transcripti
34	13	4.7	46	21	AA833049	Pinus radiata tran
35	13	4.7	59	21	AA833278	Eucalyptus grandis
36	13	4.7	67	21	AA833300	Pinus radiata tran
37	13	4.7	94	21	AA833198	Pinus radiata tran
38	13	4.7	112	21	AA833130	Pinus radiata tran
39	13	4.7	113	21	AA833025	Pinus radiata tran
40	13	4.7	114	21	AA832825	Eucalyptus grandis
41	13	4.7	130	21	AA833154	Eucalyptus grandis
42	13	4.7	148	21	AA833044	Pinus radiata tran
43	13	4.7	153	21	AA832861	Eucalyptus grandis
44	13	4.7	175	21	AA818018	Arabidopsis thalia
45	13	4.7	176	21	AA827867	Arabidopsis thalia
46	13	4.7	175	21	AA836520	Arabidopsis thalia
47	13	4.7	191	21	AA824972	Arabidopsis thalia
48	13	4.7	191	21	AA845933	Arabidopsis thalia
49	13	4.7	195	21	AA833294	Eucalyptus grandis
50	13	4.7	202	21	AA818017	Arabidopsis thalia
51	13	4.7	202	21	AA827866	Arabidopsis thalia
52	13	4.7	206	21	AA818016	Arabidopsis thalia
53	13	4.7	206	21	AA827865	Arabidopsis thalia
54	13	4.7	208	21	AA833295	Eucalyptus grandis
55	13	4.7	214	21	AA824971	Arabidopsis thalia
56	13	4.7	214	21	AA845932	Arabidopsis thalia
57	13	4.7	224	21	AA832870	Eucalyptus grandis
58	13	4.7	271	21	AA807354	Arabidopsis thalia
59	13	4.7	273	20	AAU05831	Arabidopsis thalia
60	13	4.7	307	23	AAU93066	Arabidopsis transc
61	13	4.7	309	22	AA867830	Amino acid sequenc
62	13	4.7	310	21	AA820312	Arabidopsis thalia
63	13	4.7	310	21	AA838738	Arabidopsis thalia
64	13	4.7	321	23	AAU75802	Rice morphogenesis
65	13	4.7	323	23	AAU93178	Arabidopsis transc
66	13	4.7	374	21	AAU77980	A. thaliana enviro
67	12	4.4	39	21	AA833384	Pinus radiata tran
68	12	4.4	39	21	AA833384	Pinus radiata tran
69	12	4.4	45	21	AA833371	Pinus radiata tran
70	12	4.4	46	21	AA833368	Pinus radiata tran
71	12	4.4	67	21	AA833377	Pinus radiata tran
72	12	4.4	67	21	AA833390	Pinus radiata tran
73	12	4.4	67	21	AA833391	Eucalyptus grandis
74	12	4.4	67	21	AA833407	Pinus radiata tran
75	12	4.4	67	21	AA833409	Pinus radiata tran
76	12	4.4	67	21	AA833411	Pinus radiata tran
77	12	4.4	67	21	AA833413	Pinus radiata tran
78	12	4.4	101	21	AA833214	Pinus radiata tran
79	12	4.4	101	21	AA833302	Pinus radiata tran
80	12	4.4	107	21	AA833069	Pinus radiata tran
81	12	4.4	107	21	AA833201	Pinus radiata tran
82	12	4.4	107	21	AA833275	Pinus radiata tran

83	12	4.4	107	21	AAB33289	Pinus radiata tran
84	12	4.4	120	23	ABJ10431	Myb-related transc
85	12	4.4	123	21	AAB33244	Pinus radiata tran
86	12	4.4	144	21	AAB33305	Pinus radiata tran
87	12	4.4	149	23	ABJ10421	Myb-related transc
88	12	4.4	176	21	AAB33283	Pinus radiata tran
89	12	4.4	187	21	AAB32794	Eucalyptus grandis
90	12	4.4	200	21	AAG30443	Arabidopsis thalia
91	12	4.4	205	21	AAB33293	Eucalyptus grandis
92	12	4.4	213	21	AAB32680	Eucalyptus grandis
93	12	4.4	225	21	AAB33322	Pinus radiata tran
94	12	4.4	236	21	AAG08529	Arabidopsis thalia
95	12	4.4	236	21	AAG49909	Arabidopsis thalia
96	12	4.4	236	22	AAB67842	Amino acid sequenc
97	12	4.4	246	21	AAG30442	Arabidopsis thalia
98	12	4.4	246	23	ABJ10422	Myb-related transc
99	12	4.4	253	23	ABJ10420	Myb-related transc
100	12	4.4	255	21	AAB33306	Pinus radiata tran
101	12	4.4	257	21	AAG30441	Arabidopsis thalia
102	12	4.4	268	23	ABJ10423	Myb-related transc
103	12	4.4	273	23	ABJ10416	Myb-related transc
104	12	4.4	282	21	AAG32956	Arabidopsis thalia
105	12	4.4	282	22	AAB02476	Arabidopsis thalia
106	12	4.4	282	24	AAB30047	Arabidopsis thalia
107	12	4.4	293	22	AAB82481	Cotton transcripti
108	12	4.4	295	23	ABJ10418	Myb-related transc
109	12	4.4	391	21	AAB33304	Pinus radiata tran
110	11	4.0	45	21	AAB33329	Eucalyptus grandis
111	11	4.0	61	21	AAB33339	Pinus radiata tran
112	11	4.0	62	21	AAB33366	Pinus radiata tran
113	11	4.0	63	21	AAB33393	Eucalyptus grandis
114	11	4.0	112	21	AAB33063	Pinus radiata tran
115	11	4.0	112	21	AAG25683	Arabidopsis thalia
116	11	4.0	153	21	AAG25682	Arabidopsis thalia
117	11	4.0	286	22	AAB01900	Myb-related transc
118	10	3.6	103	23	ABJ10428	Arabidopsis thalia
119	10	3.6	342	23	AAG93088	Arabidopsis thalia
120	10	3.6	371	22	AAB01913	Arabidopsis thalia
121	10	3.6	371	23	AAG92995	Arabidopsis thalia
122	9	3.3	72	21	AAB33062	Pinus radiata tran
123	9	3.3	125	21	AAG44889	Zea mays protein f
124	9	3.3	158	21	AAG44888	Zea mays protein f
125	8	2.9	44	21	AAB33352	Eucalyptus grandis
126	8	2.9	72	21	AAB33181	Eucalyptus grandis
127	8	2.9	119	21	AAB33033	Pinus radiata tran
128	8	2.9	159	21	AAB16773	Bacteriophage Dp-1
129	8	2.9	167	21	AAG26809	Zea mays protein f
130	8	2.9	281	23	AAB89612	Human polypeptide
131	8	2.9	289	22	AAG91163	C glutamicum prote
132	8	2.9	1247	16	AAR75548	MEK5 protein. Mu
133	8	2.9	1247	19	AAR56161	A mitogen-activate
134	8	2.9	1247	20	AAY33322	Mitogen ERK kinase
135	8	2.9	1247	20	AAY73555	MEK5 protein. Ma
136	8	2.9	1247	21	AAB01220	Murine MEK5. Ma
137	8	2.9	1491	23	AAB33209	Murine MEK5. Ma
138	8	2.9	1597	20	AAY42110	Murine MEK4 prote
139	8	2.9	1597	22	AAM68940	Murine MEK4-2. M
140	8	2.9	1605	22	AAG33182	Novel human secret
141	8	2.9	1607	23	AAB33210	Human mitogen-acti
142	7	2.6	49	21	AAB33350	Pinus radiata tran
143	7	2.6	50	21	AAB33350	Eucalyptus grandis
144	7	2.6	51	21	AAB33381	Pinus radiata tran
145	7	2.6	55	21	AAB33378	Pinus radiata tran
146	7	2.6	56	21	AAB33379	Pinus radiata tran
147	7	2.6	56	22	ABG10717	Novel human diagno
148	7	2.6	56	22	ABG10719	Novel human diagno
149	7	2.6	57	21	AAB33331	Eucalyptus grandis
150	7	2.6	58	21	AAB32812	Eucalyptus grandis
151	7	2.6	58	21	AAB33389	Eucalyptus grandis
152	7	2.6	58	21	AAG50021	Arabidopsis thalia
153	7	2.6	59	21	AAB33326	Eucalyptus grandis
154	7	2.6	59	21	AAB33336	Eucalyptus grandis
155	7	2.6	59	21	AAB33336	Eucalyptus grandis
156	7	2.6	156	21	AAB33359	Pinus radiata tran
157	7	2.6	157	21	AAB33367	Pinus radiata tran
158	7	2.6	158	21	AAB33333	Eucalyptus grandis
159	7	2.6	159	21	AAB33344	Eucalyptus grandis
160	7	2.6	160	21	AAB33353	Eucalyptus grandis
161	7	2.6	161	21	AAB33355	Eucalyptus grandis
162	7	2.6	162	21	AAB33369	Pinus radiata tran
163	7	2.6	163	21	AAB33370	Pinus radiata tran
164	7	2.6	164	21	AAB33376	Pinus radiata tran
165	7	2.6	165	21	AAB33385	Pinus radiata tran
166	7	2.6	166	21	AAB33387	Pinus radiata tran
167	7	2.6	167	21	AAB33400	Pinus radiata tran
168	7	2.6	168	21	AAB33402	Pinus radiata tran
169	7	2.6	169	21	AAB33406	Pinus radiata tran
170	7	2.6	170	21	AAB33408	Pinus radiata tran
171	7	2.6	171	21	AAB33410	Pinus radiata tran
172	7	2.6	172	21	AAB33412	Pinus radiata tran
173	7	2.6	173	21	AAB33412	Arabidopsis thalia
174	7	2.6	174	21	AAB33404	Pinus radiata tran
175	7	2.6	175	21	AAB33348	Eucalyptus grandis
176	7	2.6	176	21	AAB33387	Eucalyptus grandis
177	7	2.6	177	21	AAB33324	Eucalyptus grandis
178	7	2.6	178	21	AAB33339	Eucalyptus grandis
179	7	2.6	179	21	AAB33365	Pinus radiata tran
180	7	2.6	180	21	AAB33160	Eucalyptus grandis
181	7	2.6	181	21	AAG24562	Arabidopsis thalia
182	7	2.6	182	21	AAB32796	Eucalyptus grandis
183	7	2.6	183	21	ABP07394	Human ORFX protein
184	7	2.6	184	21	AAB33205	Pinus radiata tran
185	7	2.6	185	21	AAB33285	Pinus radiata tran
186	7	2.6	186	21	AAB33060	Pinus radiata tran
187	7	2.6	187	21	AAG12503	Zea mays protein f
188	7	2.6	188	21	AAG32828	Eucalyptus grandis
189	7	2.6	189	21	AAB32820	Eucalyptus grandis
190	7	2.6	190	21	AAB32756	Eucalyptus grandis
191	7	2.6	191	21	AAB33197	Pinus radiata tran
192	7	2.6	192	21	ABG24718	Novel human diagno
193	7	2.6	193	21	AAG32825	Zea mays protein f
194	7	2.6	194	21	AAB33082	Pinus radiata tran
195	7	2.6	195	21	AAB33156	Eucalyptus grandis
196	7	2.6	196	21	AAB32823	Eucalyptus grandis
197	7	2.6	197	21	AAB32822	Eucalyptus grandis
198	7	2.6	198	21	AAB33171	Eucalyptus grandis
199	7	2.6	199	21	ABJ10419	Myb-related transc
200	7	2.6	200	21	AAG26784	Zea mays protein f
201	7	2.6	201	21	AAG26727	Myb-related transc
202	7	2.6	202	21	ABJ10424	Pinus radiata tran
203	7	2.6	203	21	AAB33216	Pinus radiata tran
204	7	2.6	204	21	AAB33213	Pinus radiata tran
205	7	2.6	205	21	AAG21226	Zea mays protein f
206	7	2.6	206	21	AAB33267	Pinus radiata tran
207	7	2.6	207	21	AAB32824	Eucalyptus grandis
208	7	2.6	208	21	AAG32927	Zea mays protein f
209	7	2.6	209	21	AAB33184	Eucalyptus grandis
210	7	2.6	210	21	AAB33263	Pinus radiata tran
211	7	2.6	211	21	ABJ10407	Myb-related transc
212	7	2.6	212	21	AAG27199	Zea mays protein f
213	7	2.6	213	21	AAB33036	Pinus radiata tran
214	7	2.6	214	21	ABJ10425	Myb-related transc
215	7	2.6	215	21	AAG32926	Zea mays protein f
216	7	2.6	216	21	AAB33061	Pinus radiata tran
217	7	2.6	217	21	AAB33155	Eucalyptus grandis
218	7	2.6	218	21	AAB33260	Pinus radiata tran
219	7	2.6	219	21	AAB33251	Eucalyptus grandis
220	7	2.6	220	21	AAG27027	Zea mays protein f
221	7	2.6	221	21	AAB33196	Pinus radiata tran
222	7	2.6	222	21	AAG09858	Arabidopsis thalia
223	7	2.6	223	21	AAG31302	Arabidopsis thalia
224	7	2.6	224	21	AAG12914	Arabidopsis thalia
225	7	2.6	225	21	AAG44280	Arabidopsis thalia
226	7	2.6	226	21	AAB33206	Pinus radiata tran
227	7	2.6	227	21	AAG26781	Zea mays protein f
228	7	2.6	228	21	AAG09857	Arabidopsis thalia



375	7	2.6	469	23	AAU71802	MYB factor-1 (MYB-	448	6	2.2	57	22	AAU91744	C glutamicum prote
376	7	2.6	470	22	AAAB4558	Amino acid sequenc	449	6	2.2	57	23	ABP00255	Human ORFX protein
377	7	2.6	478	21	AAAG6852	Arabidopsis thalia	450	6	2.2	58	21	AAU23408	Arabidopsis thalia
378	7	2.6	478	21	AAAG6852	Arabidopsis thalia	451	6	2.2	58	21	ABG17983	Novel human diagno
379	7	2.6	502	23	ABBS4671	Lactococcus lactis	452	6	2.2	59	18	AAU22517	Type II topoisomer
380	7	2.6	531	21	AAAG6851	Arabidopsis thalia	453	6	2.2	59	18	AAU22518	Type II topoisomer
381	7	2.6	531	21	AAAG6851	Arabidopsis thalia	454	6	2.2	59	22	AAU53001	Protonibacterium
382	7	2.6	553	18	AAU10559	Barley gibberellin	455	6	2.2	60	21	AAU33351	Eucalyptus grandis
383	7	2.6	553	18	AAU10559	Rice gibberellin-r	456	6	2.2	60	21	AAU33351	Arabidopsis thalia
384	7	2.6	553	18	AAU10559	Rice gibberellin-r	457	6	2.2	60	21	AAU33351	Human secreted pro
385	7	2.6	683	22	ABG17032	Novel human diagno	458	6	2.2	61	21	AAU14563	Arabidopsis thalia
386	7	2.6	793	23	AAU37330	Chlamydia trachoma	459	6	2.2	61	21	AAU37330	Arabidopsis thalia
387	7	2.6	830	20	ABBO4809	LDL receptor bindi	460	6	2.2	64	21	AAU33354	Eucalyptus grandis
388	7	2.6	847	22	ABG23676	Novel human diagno	461	6	2.2	64	22	ABG26795	Novel human diagno
389	7	2.6	954	22	ABG17949	Drosophila melanog	462	6	2.2	65	23	ABP35419	Human ORFX392 prot
390	7	2.6	1013	22	ABG60789	Drosophila melanog	463	6	2.2	66	21	AAU33405	Pinus radiata tran
391	7	2.6	1245	21	AAU18244	Plasmodium falcipa	464	6	2.2	66	22	AAU60870	Protonibacterium
392	7	2.6	1316	21	AAU30505	A calcium-dependen	465	6	2.2	66	22	ABP51856	Human colon specifi
393	7	2.6	1337	21	AAU30505	A calcium-dependen	466	6	2.2	67	21	AAU33345	Eucalyptus grandis
394	7	2.6	1883	22	ABG60038	Drosophila melanog	467	6	2.2	67	21	AAU33345	Arabidopsis thalia
395	6	2.2	10	22	AAU67134	Saccharomyces cere	468	6	2.2	67	21	AAU34841	Human ORF251 prot
396	6	2.2	14	22	AAU66997	Human peptide #272	469	6	2.2	68	23	ABP33578	Protonibacterium
397	6	2.2	15	15	AAU45656	Human progesterone	470	6	2.2	69	22	AAU65870	Human liver peptid
398	6	2.2	15	16	AAU49327	Human progesterone	471	6	2.2	70	22	ABG53514	Protonibacterium
399	6	2.2	20	22	AAU51256	Oxalobacter formig	472	6	2.2	70	22	AAU62064	Peptide #6160 enco
400	6	2.2	21	23	AAU24996	TUBB1 C-terminal p	473	6	2.2	70	22	ABG38654	Peptide #5735 enco
401	6	2.2	29	23	AAU49625	Human beta-defensi	474	6	2.2	70	22	ABG23736	Human brain expre
402	6	2.2	31	18	AAU24750	Finger 2 domain of	475	6	2.2	70	22	AAU59285	Human bone marrow
403	6	2.2	34	23	AAU20615	Peptide #5 encoded	476	6	2.2	70	22	AAU71830	Peptide #5759 enco
404	6	2.2	34	23	AAU20625	Peptide #5 encoded	477	6	2.2	70	22	AAU19325	Peptide #6152 enco
405	6	2.2	35	21	AAU20274	Mouse Op-3 finger	478	6	2.2	70	22	AAU32115	Human colon specifi
406	6	2.2	35	21	AAU20274	Finger 2 subdomain	479	6	2.2	70	22	ABP51857	Human peptide enco
407	6	2.2	36	23	AAU68342	Fragment of dishev	480	6	2.2	71	21	AAU431217	Arabidopsis thalia
408	6	2.2	36	23	AAU68342	Deh peptide, DMS.	481	6	2.2	71	21	AAU431217	Protonibacterium
409	6	2.2	37	21	AAU09507	Murine Op-3 finger	482	6	2.2	71	22	AAU54255	Protonibacterium
410	6	2.2	37	22	ABG54043	Human liver peptid	483	6	2.2	73	21	ABG09383	Hepatitis GB virus
411	6	2.2	37	22	ABG39121	Peptide #6627 enco	484	6	2.2	75	22	ABG10722	Novel human diagno
412	6	2.2	37	22	ABG24018	Protein #6017 enco	485	6	2.2	76	22	ABG21320	Novel human diagno
413	6	2.2	37	22	AAU59773	Human brain expre	486	6	2.2	77	22	AAU69586	ERA binding domain
414	6	2.2	37	22	AAU72356	Human bone marrow	487	6	2.2	78	22	AAU51106	Protonibacterium
415	6	2.2	37	22	AAU19553	Peptide #5987 enco	488	6	2.2	78	22	AAU20894	Human novel foetal
416	6	2.2	37	22	AAU32614	Peptide #6651 enco	489	6	2.2	78	22	AAU21095	Human cardiovascu
417	6	2.2	37	22	ABG42172	Human peptide enco	490	6	2.2	78	23	ABG31132	D. muscipula parti
418	6	2.2	40	20	AAU27336	Group B Streptococ	491	6	2.2	78	23	ABP09114	Human ORFX protein
419	6	2.2	42	22	AAU689976	Human immune/haema	492	6	2.2	79	22	AAU91495	Human ORFX protein
420	6	2.2	43	21	AAU33386	Pinus radiata tran	493	6	2.2	79	23	ABP03105	Human ORFX protein
421	6	2.2	45	21	AAU33386	Pinus radiata tran	494	6	2.2	80	21	AAU58700	Arabidopsis thalia
422	6	2.2	45	21	AAU33386	Human lung specifi	495	6	2.2	82	22	ABG96442	Human testicular a
423	6	2.2	46	21	AAU33356	Arabidopsis thalia	496	6	2.2	82	22	AAU95911	Human reproductive
424	6	2.2	47	20	AAU00317	Human secreted pro	497	6	2.2	82	22	AAU16507	Human novel secret
425	6	2.2	47	22	ABG00326	Novel human diagno	498	6	2.2	82	23	ABG71278	Human flavoprotein
426	6	2.2	47	22	ABG22767	Novel human diagno	499	6	2.2	82	24	ABU55576	Human novel polype
427	6	2.2	49	21	AAU33984	Arabidopsis thalia	500	6	2.2	84	22	ABG96328	Human testicular a
428	6	2.2	49	21	ABG63268	Human prostate spe	501	6	2.2	84	22	AAU59974	Protonibacterium
429	6	2.2	50	22	AAU22324	Human cardiovascu	502	6	2.2	84	22	AAU22853	Human prostate can
430	6	2.2	50	22	AAU22324	A-amp protein frag	503	6	2.2	84	22	AAU95797	Human reproductive
431	6	2.2	51	22	ABG31968	Novel human diagno	504	6	2.2	84	22	AAU96159	Human reproductive
432	6	2.2	52	23	ABP32185	Human ORF1158 prot	505	6	2.2	84	22	AAU20628	Human secreted pro
433	6	2.2	53	22	AAU53204	Protonibacterium	506	6	2.2	85	21	AAU15240	Arabidopsis thalia
434	6	2.2	53	22	AAU65580	Protonibacterium	507	6	2.2	86	22	AAU91998	Human digestive sy
435	6	2.2	55	22	ABG48908	Human liver peptid	508	6	2.2	86	23	ABP41258	Human ovarian anti
436	6	2.2	55	22	ABG82906	Peptide #1587 enco	509	6	2.2	86	23	ABG51355	Human ORF61 protei
437	6	2.2	55	22	ABG34081	Peptide #1587 enco	510	6	2.2	87	22	ABG14369	Novel human diagno
438	6	2.2	55	22	ABG19517	Protein #1516 enco	511	6	2.2	88	21	AAU25632	Arabidopsis thalia
439	6	2.2	55	22	AAU67484	Human brain expre	512	6	2.2	88	21	AAU61192	Nuclear transpor
440	6	2.2	55	22	AAU67245	Human bone marrow	513	6	2.2	89	22	AAU61244	Protonibacterium
441	6	2.2	55	22	AAU15090	Peptide #1524 enco	514	6	2.2	90	23	ABP03386	Human ORFX protein
442	6	2.2	55	22	AAU27539	Peptide #1576 enco	515	6	2.2	91	22	AAU74654	Human colon cancer
443	6	2.2	55	22	AAU02826	Peptide #1508 enco	516	6	2.2	93	21	AAU15239	Arabidopsis thalia
444	6	2.2	55	22	ABG36896	Human peptide enco	517	6	2.2	93	22	AAU54461	Protonibacterium
445	6	2.2	56	23	ABG02992	Novel human diagno	518	6	2.2	93	22	AAU01737	Human polypeptide
446	6	2.2	57	20	AAU74232	Human prostate tum	519	6	2.2	94	22	ABG13591	Novel human diagno
447	6	2.2	57	21	AAU51747	Arabidopsis thalia	520	6	2.2	95	21	AAU10041	LCMW unknown prote

521	6	2.2	95	22	AA00823	Human polypeptide	594	6	2.2	145	23	ABB48400	Listeria monocytog
522	6	2.2	97	20	AAE60346	Human normal bladd	595	6	2.2	145	24	ABP76254	Human GENSER prote
523	6	2.2	97	21	AAE29343	Arabidopsis thalia	596	6	2.2	147	22	ABB64524	Drosophila melanog
524	6	2.2	97	22	AAU30967	Novel human secret	597	6	2.2	148	22	ABG29364	Novel human diagno
525	6	2.2	101	22	ABG26051	Novel human diagno	598	6	2.2	148	22	AAO00081	Human polypeptide
526	6	2.2	101	22	ABG4985	Human protein sequ	599	6	2.2	150	23	ABB53480	Lactococcus lactis
527	6	2.2	102	21	AAAB09536	Murine OP-3, SEQ I	600	6	2.2	150	24	ABU02510	S. pneumoniae type
528	6	2.2	102	21	AAAB02802	Mouse OP-3 amino a	601	6	2.2	152	21	AAAB63217	Gene 42 human secr
529	6	2.2	102	21	AAAY92571	OP-3 finger-1-heel	602	6	2.2	152	21	AAAB63218	Human secreted pro
530	6	2.2	103	20	AAAY5812	Chlamydia pneumoni	603	6	2.2	152	21	AAAC24037	Arabidopsis thalia
531	6	2.2	105	23	ABP00976	Human ORFX protein	604	6	2.2	153	22	AAU40542	Propionibacterium
532	6	2.2	105	23	ABP04209	Human ORFX protein	605	6	2.2	153	22	ABG07537	Novel human diagno
533	6	2.2	107	22	ABG12930	Novel human diagno	606	6	2.2	153	22	ABG20018	Novel human diagno
534	6	2.2	107	23	ABP05745	Human ORFX protein	607	6	2.2	153	23	ABG97433	S. carnositatus
535	6	2.2	108	21	AAAB21190	Exo14 partial prot	608	6	2.2	155	21	AAAB42104	Human ORFX ORF168
536	6	2.2	108	22	AAAW76675	Human protein SEQ	609	6	2.2	156	21	AAAC27380	Arabidopsis thalia
537	6	2.2	110	21	AAAG10027	Arabidopsis thalia	609	6	2.2	156	21	AAAC27380	Arabidopsis thalia
538	6	2.2	110	22	ABG48162	Human liver peptid	610	6	2.2	158	19	AAW71686	E. coli Repa prote
539	6	2.2	110	22	ABAB28138	Human peptid #789	611	6	2.2	158	21	AAW77570	Human cytoskeletal
540	6	2.2	110	22	ABAB33313	Peptide #819 encod	612	6	2.2	159	22	AAU17299	Novel signal trans
541	6	2.2	110	22	ABAB18773	Protein #772 encod	613	6	2.2	160	14	AAAR33386	Max protein. Homo
542	6	2.2	110	22	AAAS4102	Human brain expres	614	6	2.2	161	22	AAAB69531	DRAP protein. Dro
543	6	2.2	110	22	AAAG66492	Human bone marrow	615	6	2.2	163	22	ABG12960	Novel human diagno
544	6	2.2	110	22	AAAM14365	Peptide #799 encod	616	6	2.2	165	21	AAAG25631	Arabidopsis thalia
545	6	2.2	110	22	AAAM26778	Peptide #815 encod	617	6	2.2	165	22	AAU67372	Propionibacterium
546	6	2.2	110	22	AAAM02094	Peptide #776 encod	618	6	2.2	165	23	ABP64737	Human protein SEQ
547	6	2.2	110	23	ABG36144	Human peptid encod	619	6	2.2	165	24	ABP60632	Nebulin 18.15. Un
548	6	2.2	111	21	AAAG0639	Zea mays protein f	620	6	2.2	168	22	AAU629848	Propionibacterium
549	6	2.2	112	21	AAAG43521	Propionibacterium	621	6	2.2	168	22	AAAG91968	C glutamicum prote
550	6	2.2	112	22	ABAB10981	Human GPCR hRUP5 h	622	6	2.2	168	23	ABB55243	Lactococcus lactis
551	6	2.2	112	22	ABP03621	Human ORFX protein	623	6	2.2	169	18	AAWL6321	Murine ARF-p19, a
552	6	2.2	113	21	AAAB32818	Bucalyptus grandis	624	6	2.2	169	21	AAW79143	Mouse ARF-p19 cell
553	6	2.2	113	22	AAAG61717	Arabidopsis thalia	625	6	2.2	169	22	AAU29848	Novel human secret
554	6	2.2	113	22	AAU42676	Propionibacterium	626	6	2.2	169	22	AAAM39453	Human polypeptide
555	6	2.2	113	23	ABP01998	Human ORFX protein	627	6	2.2	169	23	AAE25910	Mouse ARF-p19 prot
556	6	2.2	115	21	AAAG12915	Arabidopsis thalia	628	6	2.2	169	23	AAO15373	Mouse Arf protein.
557	6	2.2	116	21	AAAB33072	Pinus radiata tran	629	6	2.2	171	24	ABP81647	Streptococcus pneu
558	6	2.2	116	22	AAUS3832	Propionibacterium	630	6	2.2	172	22	ABAB1433	Human intelectin h
559	6	2.2	116	23	ABP01081	Human ORFX protein	631	6	2.2	172	23	ABP25691	Streptococcus poly
560	6	2.2	120	21	AAAG60639	Arabidopsis thalia	632	6	2.2	175	21	AAAG27379	Arabidopsis thalia
561	6	2.2	121	21	AAAG7381	Arabidopsis thalia	633	6	2.2	176	21	AAAB27867	Sequence homology
562	6	2.2	121	21	AAAG34840	Arabidopsis thalia	634	6	2.2	177	23	AAAB47739	Listeria monocytog
563	6	2.2	121	21	AAAG44064	Arabidopsis thalia	635	6	2.2	178	22	AAAB64632	Human secreted pro
564	6	2.2	121	23	ABAB49279	Listeria monocytog	636	6	2.2	180	21	AAAG29341	Arabidopsis thalia
565	6	2.2	122	22	AAUS9871	Propionibacterium	637	6	2.2	180	21	AAAG00814	Human secreted pro
566	6	2.2	123	22	AAU20123	Human DNA repair a	638	6	2.2	180	22	AAU39309	Propionibacterium
567	6	2.2	123	23	ABG91372	Novel human DNA re	639	6	2.2	181	22	ABG06860	Novel human diagno
568	6	2.2	124	21	AAAG29342	Arabidopsis thalia	640	6	2.2	183	22	AAAB64631	Human secreted pro
569	6	2.2	124	23	ABP03440	Human ORFX protein	641	6	2.2	186	22	AAU01618	Human secreted pro
570	6	2.2	125	21	AAAG02748	Human secreted pro	642	6	2.2	187	23	ABP33465	Human oxidase-like
571	6	2.2	125	22	AAO05636	Human polypeptide	643	6	2.2	191	22	ABG08039	Novel human diagno
572	6	2.2	126	20	AAI25374	Human b2-HPLS2OW c	644	6	2.2	191	22	AAE05402	Truncated from of
573	6	2.2	126	21	AAAG11567	Arabidopsis thalia	645	6	2.2	192	24	ABR41162	Human DITP trans
574	6	2.2	126	21	AAAG11567	Arabidopsis thalia	646	6	2.2	193	24	ABP78438	N. gonorrhoeae ami
575	6	2.2	127	20	AAAY30400	Human clone p12616	647	6	2.2	194	24	AAAG91250	C glutamicum prote
576	6	2.2	127	22	AAAM64385	CASB414 polypeptid	648	6	2.2	194	24	ABU51202	Helicobacter pylor
577	6	2.2	127	22	AAAM77297	Human brain expres	649	6	2.2	194	24	ABP81341	Streptococcus pneu
578	6	2.2	127	22	ABG46218	Human bone marrow	650	6	2.2	194	24	ABU01001	Streptococcus pneu
579	6	2.2	127	23	ABP33457	Human peptid encod	651	6	2.2	196	20	AAV76622	S. pneumoniae type
580	6	2.2	128	22	AAU40432	Human ORF2430 prot	652	6	2.2	196	20	AAV35786	Human aviarian tumo
581	6	2.2	130	23	ABP43829	Propionibacterium	653	6	2.2	196	20	AAV02827	Amino acid sequenc
582	6	2.2	130	23	ABP43829	T-cell receptor al	654	6	2.2	196	20	ABBS3881	Lactococcus lactis
583	6	2.2	132	21	AAAG40836	Zea mays protein f	655	6	2.2	198	22	ABR71919	Drosophila melanog
584	6	2.2	135	21	AAAG18449	Propionibacterium	656	6	2.2	199	22	AAAG90817	C glutamicum prote
585	6	2.2	135	22	AAUS2049	Propionibacterium	657	6	2.2	201	22	AAW41139	Human polypeptide
586	6	2.2	137	22	AAU43202	Human TGFbeta prot	658	6	2.2	204	21	AAAG31516	Arabidopsis thalia
587	6	2.2	139	23	AAAM51948	Human TGFbeta prot	659	6	2.2	204	24	AAE35125	Woody plant hybrid
588	6	2.2	140	21	AAAG26134	Zea mays protein f	660	6	2.2	205	22	ABG13705	Arabidopsis thalia
589	6	2.2	140	22	ABG19382	Novel human diagno	661	6	2.2	206	21	AAAG27684	Novel human diagno
590	6	2.2	142	22	AAAG15690	Human immune/haema	662	6	2.2	206	21	AAV92247	Human cancer assoc
591	6	2.2	143	22	AAU65348	Propionibacterium	663	6	2.2	207	23	AAE25364	Streptococcus pyog
592	6	2.2	143	22	ABG00273	Novel human diagno	664	6	2.2	207	23	AAE25365	S. pyogenes pyroge
593	6	2.2	143	23	AAE25363	Streptococcus pyog	665	6	2.2	207	23	AAE25367	S. pyogenes pyroge
	6	2.2	144	22	AAAB79470	Corynebacterium gl	666	6	2.2	207	23	AAE25368	S. pyogenes pyroge

667	6	2.2	207	23	AAE25369	S. pyogenes pyroge	740	6	2.2	252	21	AAV73366	HTM clone 2918375
668	6	2.2	207	23	AAE25370	S. pyogenes pyroge	741	6	2.2	252	23	ABB48700	Listeria monocytog
669	6	2.2	207	23	AAE25371	S. pyogenes pyroge	742	6	2.2	253	21	AA606682	Arabidopsis thalia
670	6	2.2	207	23	AAE25372	S. pyogenes pyroge	743	6	2.2	253	23	ABU52188	Helicobacter pylori
671	6	2.2	207	23	AAE25373	S. pyogenes pyroge	744	6	2.2	255	22	ABG26755	Novel human diagno
672	6	2.2	207	23	AAE25374	S. pyogenes pyroge	745	6	2.2	256	20	AAJ37115	Protein involved i
673	6	2.2	207	23	AAE25395	S. pyogenes pyroge	746	6	2.2	257	21	AA606681	Arabidopsis thalia
674	6	2.2	208	12	AAE25395	S. pyogenes pyroge	747	6	2.2	257	21	AA606681	Arabidopsis thalia
675	6	2.2	208	14	AAE25395	Streptococcus pyro	748	6	2.2	257	21	AA606681	Arabidopsis thalia
676	6	2.2	208	14	AAE25395	Streptococcus pyro	749	6	2.2	257	21	AA606681	Arabidopsis thalia
677	6	2.2	208	14	AAE25395	Streptococcus pyro	750	6	2.2	257	21	AA606681	Arabidopsis thalia
678	6	2.2	208	14	AAE25395	Streptococcus pyro	751	6	2.2	257	21	AA606681	Arabidopsis thalia
679	6	2.2	208	14	AAE25395	Streptococcus pyro	752	6	2.2	259	21	AA606681	Arabidopsis thalia
680	6	2.2	209	22	AAU43183	Streptococcus pyro	753	6	2.2	261	21	AA629334	Arabidopsis thalia
681	6	2.2	211	21	AAU43183	Streptococcus pyro	754	6	2.2	261	21	AA629334	Arabidopsis thalia
682	6	2.2	211	21	AAU43183	Streptococcus pyro	755	6	2.2	262	18	AAW38194	Arabidopsis thalia
683	6	2.2	212	23	ABP73477	Arabidopsis thalia	756	6	2.2	262	18	AAW38194	Arabidopsis thalia
684	6	2.2	215	23	ABP73477	Arabidopsis thalia	757	6	2.2	263	20	AAW38194	Arabidopsis thalia
685	6	2.2	217	21	AAU43183	Arabidopsis thalia	758	6	2.2	263	20	AAW38194	Arabidopsis thalia
686	6	2.2	218	20	AAU43183	Arabidopsis thalia	759	6	2.2	263	20	AAW38194	Arabidopsis thalia
687	6	2.2	218	20	AAU43183	Arabidopsis thalia	760	6	2.2	263	20	AAW38194	Arabidopsis thalia
688	6	2.2	219	22	AAU43183	Arabidopsis thalia	761	6	2.2	263	20	AAW38194	Arabidopsis thalia
689	6	2.2	220	21	AAU43183	Arabidopsis thalia	762	6	2.2	263	20	AAW38194	Arabidopsis thalia
690	6	2.2	220	21	AAU43183	Arabidopsis thalia	763	6	2.2	263	20	AAW38194	Arabidopsis thalia
691	6	2.2	220	21	AAU43183	Arabidopsis thalia	764	6	2.2	263	20	AAW38194	Arabidopsis thalia
692	6	2.2	220	21	AAU43183	Arabidopsis thalia	765	6	2.2	263	20	AAW38194	Arabidopsis thalia
693	6	2.2	220	21	AAU43183	Arabidopsis thalia	766	6	2.2	263	20	AAW38194	Arabidopsis thalia
694	6	2.2	222	21	AAU43183	Arabidopsis thalia	767	6	2.2	263	20	AAW38194	Arabidopsis thalia
695	6	2.2	222	21	AAU43183	Arabidopsis thalia	768	6	2.2	263	20	AAW38194	Arabidopsis thalia
696	6	2.2	222	21	AAU43183	Arabidopsis thalia	769	6	2.2	263	20	AAW38194	Arabidopsis thalia
697	6	2.2	225	22	AAU43183	Arabidopsis thalia	770	6	2.2	263	20	AAW38194	Arabidopsis thalia
698	6	2.2	225	22	AAU43183	Arabidopsis thalia	771	6	2.2	263	20	AAW38194	Arabidopsis thalia
699	6	2.2	225	22	AAU43183	Arabidopsis thalia	772	6	2.2	263	20	AAW38194	Arabidopsis thalia
700	6	2.2	227	21	AAU43183	Arabidopsis thalia	773	6	2.2	263	20	AAW38194	Arabidopsis thalia
701	6	2.2	227	21	AAU43183	Arabidopsis thalia	774	6	2.2	264	22	AAW38194	Arabidopsis thalia
702	6	2.2	228	21	AAU43183	Arabidopsis thalia	775	6	2.2	266	22	AAW38194	Arabidopsis thalia
703	6	2.2	235	19	AAU43183	Arabidopsis thalia	776	6	2.2	267	22	AAW38194	Arabidopsis thalia
704	6	2.2	235	19	AAU43183	Arabidopsis thalia	777	6	2.2	269	21	AAW38194	Arabidopsis thalia
705	6	2.2	235	19	AAU43183	Arabidopsis thalia	778	6	2.2	269	21	AAW38194	Arabidopsis thalia
706	6	2.2	235	19	AAU43183	Arabidopsis thalia	779	6	2.2	269	21	AAW38194	Arabidopsis thalia
707	6	2.2	235	19	AAU43183	Arabidopsis thalia	780	6	2.2	270	21	AAW38194	Arabidopsis thalia
708	6	2.2	235	19	AAU43183	Arabidopsis thalia	781	6	2.2	271	19	AAW38194	Arabidopsis thalia
709	6	2.2	235	19	AAU43183	Arabidopsis thalia	782	6	2.2	271	19	AAW38194	Arabidopsis thalia
710	6	2.2	236	20	AAU43183	Arabidopsis thalia	783	6	2.2	271	19	AAW38194	Arabidopsis thalia
711	6	2.2	236	20	AAU43183	Arabidopsis thalia	784	6	2.2	271	19	AAW38194	Arabidopsis thalia
712	6	2.2	236	20	AAU43183	Arabidopsis thalia	785	6	2.2	271	19	AAW38194	Arabidopsis thalia
713	6	2.2	237	21	AAU43183	Arabidopsis thalia	786	6	2.2	273	18	AAW38194	Arabidopsis thalia
714	6	2.2	238	21	AAU43183	Arabidopsis thalia	787	6	2.2	273	18	AAW38194	Arabidopsis thalia
715	6	2.2	240	22	AAU43183	Arabidopsis thalia	788	6	2.2	273	18	AAW38194	Arabidopsis thalia
716	6	2.2	240	22	AAU43183	Arabidopsis thalia	789	6	2.2	274	23	AAW38194	Arabidopsis thalia
717	6	2.2	242	22	AAU43183	Arabidopsis thalia	790	6	2.2	276	20	AAW38194	Arabidopsis thalia
718	6	2.2	244	21	AAU43183	Arabidopsis thalia	791	6	2.2	277	18	AAW38194	Arabidopsis thalia
719	6	2.2	246	21	AAU43183	Arabidopsis thalia	792	6	2.2	278	22	AAW38194	Arabidopsis thalia
720	6	2.2	247	21	AAU43183	Arabidopsis thalia	793	6	2.2	278	22	AAW38194	Arabidopsis thalia
721	6	2.2	247	21	AAU43183	Arabidopsis thalia	794	6	2.2	278	22	AAW38194	Arabidopsis thalia
722	6	2.2	247	21	AAU43183	Arabidopsis thalia	795	6	2.2	278	22	AAW38194	Arabidopsis thalia
723	6	2.2	248	19	AAU43183	Arabidopsis thalia	796	6	2.2	280	21	AAW38194	Arabidopsis thalia
724	6	2.2	248	19	AAU43183	Arabidopsis thalia	797	6	2.2	280	21	AAW38194	Arabidopsis thalia
725	6	2.2	249	18	AAU43183	Arabidopsis thalia	798	6	2.2	280	21	AAW38194	Arabidopsis thalia
726	6	2.2	249	18	AAU43183	Arabidopsis thalia	799	6	2.2	281	21	AAW38194	Arabidopsis thalia
727	6	2.2	249	18	AAU43183	Arabidopsis thalia	800	6	2.2	281	21	AAW38194	Arabidopsis thalia
728	6	2.2	249	21	AAU43183	Arabidopsis thalia	801	6	2.2	283	22	AAW38194	Arabidopsis thalia
729	6	2.2	249	21	AAU43183	Arabidopsis thalia	802	6	2.2	283	22	AAW38194	Arabidopsis thalia
730	6	2.2	249	21	AAU43183	Arabidopsis thalia	803	6	2.2	284	22	AAW38194	Arabidopsis thalia
731	6	2.2	249	21	AAU43183	Arabidopsis thalia	804	6	2.2	285	21	AAW38194	Arabidopsis thalia
732	6	2.2	250	19	AAU43183	Arabidopsis thalia	805	6	2.2	285	21	AAW38194	Arabidopsis thalia
733	6	2.2	250	19	AAU43183	Arabidopsis thalia	806	6	2.2	285	21	AAW38194	Arabidopsis thalia
734	6	2.2	250	19	AAU43183	Arabidopsis thalia	807	6	2.2	286	23	AAW38194	Arabidopsis thalia
735	6	2.2	250	20	AAU43183	Arabidopsis thalia	808	6	2.2	286	23	AAW38194	Arabidopsis thalia
736	6	2.2	250	20	AAU43183	Arabidopsis thalia	809	6	2.2	287	23	AAW38194	Arabidopsis thalia
737	6	2.2	250	22	AAU43183	Arabidopsis thalia	810	6	2.2	289	22	AAW38194	Arabidopsis thalia
738	6	2.2	250	23	AAU43183	Arabidopsis thalia	811	6	2.2	289	22	AAW38194	Arabidopsis thalia
739	6	2.2	250	24	AAU43183	Arabidopsis thalia	812	6	2.2	292	21	AAW38194	Arabidopsis thalia

813	6	2.2	293	21	AA636986	Arabidopsis thalia	886	6	2.2	318	23	AAU95699	Human olfactory an
814	6	2.2	293	22	ABG05968	Novel human diagno	887	6	2.2	318	24	ABU11220	Human G-protein co
815	6	2.2	293	22	AA682724	S. epidermidis ope	888	6	2.2	319	21	AA655058	Arabidopsis thalia
816	6	2.2	294	20	AA932887	Human mucilage pro	889	6	2.2	319	23	AAU76435	Novel desulphurisi
817	6	2.2	294	21	AA608152	Arabidopsis thalia	890	6	2.2	320	21	AA660193	Arabidopsis thalia
818	6	2.2	294	21	AA638782	Arabidopsis thalia	891	6	2.2	321	21	AA634413	Arabidopsis thalia
819	6	2.2	294	22	AA664475	Gene 8 human secre	892	6	2.2	322	21	AA625329	Arabidopsis thalia
820	6	2.2	295	19	AA684943	A truncated papill	893	6	2.2	323	21	AA628149	Arabidopsis thalia
821	6	2.2	295	20	AA631889	Corn cyclin A part	894	6	2.2	324	21	AA682096	S. epidermidis ope
822	6	2.2	295	23	ABG76833	Human G-protein co	895	6	2.2	326	21	AA657027	Human prostate can
823	6	2.2	296	22	AA692555	Human protein sequ	896	6	2.2	327	22	AA648010	Rat OR2R polypepti
824	6	2.2	298	21	AA614560	Arabidopsis thalia	897	6	2.2	328	24	ABP76728	Streptomyces virid
825	6	2.2	299	23	ABP39569	Staphylococcus epi	898	6	2.2	329	24	ABP98517	YK1018w zinc bindi
826	6	2.2	302	22	ABG07538	Novel human diagno	899	6	2.2	330	21	AA606262	Arabidopsis thalia
827	6	2.2	304	13	AA625291	Acetyl xylan ester	900	6	2.2	330	21	AA658896	Arabidopsis thalia
828	6	2.2	305	18	AAW34217	Streptomyces oxido	901	6	2.2	330	22	AA660354	Mouse atonal homol
829	6	2.2	305	19	AAW55818	Streptomyces roseo	902	6	2.2	330	22	AA660363	Mouse atonal homol
830	6	2.2	306	22	ABG10724	Novel human diagno	903	6	2.2	330	22	AA660370	Mouse atonal homol
831	6	2.2	306	24	AAE34811	Bovine papillomavi	904	6	2.2	335	23	ABP92283	Herbicidially activ
832	6	2.2	309	23	ABE49453	Listeria monocytog	905	6	2.2	335	24	ABP81239	Arabidopsis thalia
833	6	2.2	310	22	AA696075	Putative P. abyssi	906	6	2.2	336	22	AA662450	Human membrane bin
834	6	2.2	310	22	AA682321	S. epidermidis ope	907	6	2.2	336	22	AA679137	Corynebacterium gl
835	6	2.2	312	21	AA641496	Arabidopsis thalia	908	6	2.2	336	22	AA679487	Corynebacterium gl
836	6	2.2	313	20	AA434039	CAS414 polypeptid	909	6	2.2	336	23	ABG69364	Babesia amino acid
837	6	2.2	313	21	AA471905	FVB mouse TH2AF1	910	6	2.2	339	23	ABP69274	Human polypeptide
838	6	2.2	313	21	AA471906	TG5 mouse TH2AF1	911	6	2.2	341	19	AAW56423	Trypophanyl crna
839	6	2.2	313	21	AA471907	DBA2 mouse TH2AF1	912	6	2.2	341	22	AAU37851	Streptococcus pneu
840	6	2.2	313	21	AA471908	C57BL/6 mouse TH2A	913	6	2.2	341	23	ABP26964	Streptococcus poly
841	6	2.2	313	21	AA471909	Human TH2AF1 isocy	914	6	2.2	341	23	ABP26965	Streptococcus poly
842	6	2.2	313	21	AA494858	Human protein clon	915	6	2.2	341	23	ABP26966	Streptococcus poly
843	6	2.2	313	21	AA494858	Membrane-bound pro	916	6	2.2	341	24	ABU02760	Lactococcus lactis
844	6	2.2	313	22	AAU29170	Human PRO polypept	917	6	2.2	342	21	AA641495	S. pneumoniae type
845	6	2.2	313	22	AA689970	C glutamicum prote	918	6	2.2	344	23	ABP41961	Arabidopsis thalia
846	6	2.2	313	22	AA687569	Human PRO1270. Ho	919	6	2.2	347	18	AAW20595	Human ovarian anti
847	6	2.2	313	22	AA687748	Human T2R18 amino	920	6	2.2	348	22	AAW41060	H. pylori cytoplasm
848	6	2.2	313	22	AA471965	Human calcium-depe	921	6	2.2	350	21	AA614334	Human polypeptide
849	6	2.2	313	22	AA471962	Human calicum-depe	922	6	2.2	350	21	AA629759	Arabidopsis thalia
850	6	2.2	313	22	AA665283	Human PRO1270 (UNQ	923	6	2.2	350	22	ABG27189	Arabidopsis thalia
851	6	2.2	313	23	ABG58984	Human secreted/citra	924	6	2.2	351	22	AAW79701	Novel human diagno
852	6	2.2	313	24	ABU71258	Human PRO1270 prot	925	6	2.2	351	23	ABP92633	Human protein SRO
853	6	2.2	313	24	ABU71549	Human secreted pol	926	6	2.2	352	18	AAW30300	Herbicidially activ
854	6	2.2	313	24	ABU71995	Novel human secret	927	6	2.2	353	21	AA636439	Cobalt-t-transportin
855	6	2.2	313	24	ABU72152	Human PRO polypept	928	6	2.2	353	21	AA658895	Arabidopsis thalia
856	6	2.2	313	24	ABU65715	Human secreted/citra	929	6	2.2	355	21	AA628148	Arabidopsis thalia
857	6	2.2	313	24	ABU66048	Novel human secret	930	6	2.2	355	23	AAO22175	Arabidopsis thalia
858	6	2.2	313	24	ABU67552	Human secreted/citra	931	6	2.2	356	23	ABP55105	Ranoplanin biosynt
859	6	2.2	313	24	ABU65410	Human PRO polypept	932	6	2.2	357	22	AAU45695	Lactococcus lactis
860	6	2.2	313	24	ABU59176	Human PRO polypept	933	6	2.2	357	22	AAU18345	Lactococcus lactis
861	6	2.2	313	24	ABU59323	Novel human secret	934	6	2.2	361	23	ABP55078	Human endocrine po
862	6	2.2	313	24	ABU59472	Human secreted/citra	935	6	2.2	362	23	AAU80013	Lactococcus lactis
863	6	2.2	313	24	ABU60607	Novel human secret	936	6	2.2	362	23	AAU79807	Shinapyl alcohol de
864	6	2.2	313	24	ABU58098	Human PRO polypept	937	6	2.2	363	20	AAV37863	Shinapyl alcohol de
865	6	2.2	313	24	ABU58546	Human PRO polypept	938	6	2.2	365	23	ABP09748	Amino acid sequenc
866	6	2.2	313	24	ABU59029	Human secreted/citra	939	6	2.2	365	23	ABP48295	Amino acid sequenc
867	6	2.2	313	24	ABU56082	Human secreted/citra	940	6	2.2	365	24	ABP58341	Listeria monocytog
868	6	2.2	313	24	ABU57077	Human PRO polypept	941	6	2.2	368	21	AA613590	Human cell growth,
869	6	2.2	313	24	ABU13989	Human PRO1270 poly	942	6	2.2	369	21	AA623924	Streptomyces globi
870	6	2.2	313	24	ABU10656	Human secreted/citra	943	6	2.2	370	21	AA623465	Arabidopsis thalia
871	6	2.2	313	24	ABU10944	Human PRO polypept	944	6	2.2	370	21	AA626531	Arabidopsis thalia
872	6	2.2	314	20	AA625409	Human secreted pro	945	6	2.2	370	21	AA644397	Arabidopsis thalia
873	6	2.2	314	21	AA625409	Pinus radiata cell	946	6	2.2	370	21	AAV74818	Neisseria gonorrhoe
874	6	2.2	314	23	ABP43649	Clone MGC:20469.	947	6	2.2	370	21	AAV74819	Neisseria meningit
875	6	2.2	314	23	ABP37905	NOX protein sequ	948	6	2.2	370	21	AAV74820	Neisseria meningit
876	6	2.2	315	21	AA658887	Arabidopsis thalia	949	6	2.2	371	21	AA625464	Arabidopsis thalia
877	6	2.2	315	23	ABP39296	Staphylococcus epi	950	6	2.2	371	21	AA626530	Arabidopsis thalia
878	6	2.2	316	22	AA672297	Human olfactory re	951	6	2.2	371	21	AA629758	Arabidopsis thalia
879	6	2.2	316	22	AA672415	Human OR-like poly	952	6	2.2	372	21	AA614333	Arabidopsis thalia
880	6	2.2	317	21	AA617319	Arabidopsis thalia	953	6	2.2	373	21	AA630101	Arabidopsis thalia
881	6	2.2	317	22	AA672837	Human olfactory re	954	6	2.2	374	22	AA699728	Streptomyces sp. C
882	6	2.2	318	22	AA671540	Human olfactory re	955	6	2.2	374	22	ABU51613	Helicobacter pylor
883	6	2.2	318	22	AA672395	Human OR-like poly	956	6	2.2	375	22	ABG12617	Novel human diagno
884	6	2.2	318	23	ABP95657	Human GPCR polypep	957	6	2.2	376	19	AAW81744	M. tuberculosis im
885	6	2.2	318	23	AA671168	Human GPCR protei	958	6	2.2	376	19	AAW64377	Mycobacterium tube



959	6	2.2	376	20	AAV39174	M. tuberculosis an
960	6	2.2	376	20	AAV39031	M. tuberculosis re
961	6	2.2	376	20	AAV33722	Streptomyces albid
962	6	2.2	376	22	AAV33553	Amino acid sequenc
963	6	2.2	378	21	AAV34557	Arabidopsis thalia
964	6	2.2	379	22	AAV34557	Human brain expres
965	6	2.2	379	22	AAV34557	S. epidermidis ope
966	6	2.2	380	22	ABG05463	Novel human diagno
967	6	2.2	380	22	ABG10100	Human secreted pro
968	6	2.2	382	21	AAV34939	Laccococcus lactis
969	6	2.2	383	15	AAV32083	Propionibacterium
970	6	2.2	383	22	AAV33454	Human AFP protein
971	6	2.2	383	22	AAV31258	Sequence encoded b
972	6	2.2	385	15	AAV31060	Mycobacterium tube
973	6	2.2	385	22	AAV31097	Novel human diagno
974	6	2.2	386	23	ABG03421	Human ovarian anti
975	6	2.2	387	21	AAV31651	Arabidopsis thalia
976	6	2.2	387	21	AAV325463	Arabidopsis thalia
977	6	2.2	387	21	AAV325463	Arabidopsis thalia
978	6	2.2	387	22	ABG26529	Novel human diagno
979	6	2.2	387	22	AAV315104	Human AFP protein
980	6	2.2	388	21	AAV31305	Arabidopsis thalia
981	6	2.2	389	21	AAV34396	Arabidopsis thalia
982	6	2.2	391	21	AAV34396	Amino acid sequenc
983	6	2.2	391	21	AAV34396	Arabidopsis thalia
984	6	2.2	392	22	ABG07531	Novel human diagno
985	6	2.2	392	22	ABG12279	Protein regulating
986	6	2.2	392	22	AAV31963	Novel human diagno
987	6	2.2	392	22	AAV31963	Arabidopsis thalia
988	6	2.2	392	22	AAV31963	Abesic acid resp
989	6	2.2	392	22	AAV31963	A. thaliana abscis
990	6	2.2	392	22	AAV31963	Abesic acid resp
991	6	2.2	392	22	AAV31963	Abesic acid resp
992	6	2.2	392	22	AAV31963	Abesic acid resp
993	6	2.2	392	22	AAV31963	Abesic acid resp
994	6	2.2	392	22	AAV31963	Abesic acid resp
995	6	2.2	392	22	AAV31963	Abesic acid resp
996	6	2.2	392	22	AAV31963	Abesic acid resp
997	6	2.2	392	22	AAV31963	Abesic acid resp
998	6	2.2	392	22	AAV31963	Abesic acid resp
999	6	2.2	392	22	AAV31963	Abesic acid resp
1000	6	2.2	392	22	AAV31963	Abesic acid resp

## ALIGNMENTS

RESULT 1:  
ID ABB81626 standard; Protein; 274 AA.

AC ABB81626;  
XX  
XX  
XX 24-SEP-2002 (first entry)  
XX  
XX Tomato anthocyanin 1 (ANT1) protein SEQ ID NO.2.  
XX  
XX Tomato; ANT1; anthocyanin 1; Lycopersicon esculentum; plant; colour.  
XX  
XX Lycopersicon esculentum.  
XX  
XX WO200255658-A2.  
XX  
XX PD 18-JUL-2002.  
XX  
XX PF 29-OCT-2001; 2001WO-US0638.  
XX  
XX PR 30-OCT-2000; 2000US-244685P.  
XX  
XX PA (EXEL-) EXELIXIS PLANT SCI INC.  
XX  
XX PI Connors K, Mathews HV, Liu A;  
XX

DR WPI; 2002-557819/59.  
DR N-PSDB; ABQ73046.  
XX  
XX New isolated polynucleotide derived from tomato, useful for producing  
PT an Anthocyanin 1 phenotype in plants, particularly for modifying e.g.  
PT leaf color, flower color or fruit color in plants  
PT  
XX Claim 5; Page 39-41; 53pp; English.  
XX  
XX The present sequence represents tomato anthocyanin 1 (ANT1). The ANT1  
CC polynucleotide can be used for modifying the expression of a native  
CC plant gene, particularly for producing an anthocyanin 1 phenotype in  
CC plants, which is responsible for many red and blue colours in plants.  
CC The polynucleotide is useful for modifying e.g. leaf colour, flower  
CC colour or fruit colour in plants.  
XX  
XX Sequence 274 AA:

Query Match 100.0%; Score 274; DB 23; Length 274;  
Best local similarity 100.0%; Pred. No. 2.4e-262;  
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNSTSSSLGVRKGSWTDEEDFLRKCIDKYGKWHVPIRAGLNRCKRCKRLRANTY 60  
DB 1 MNSTSSSLGVRKGSWTDEEDFLRKCIDKYGKWHVPIRAGLNRCKRCKRLRANTY 60  
QY 61 RPIIKSGDEPDEVDILRLKLLGNRSLIAGRLPGRTANDVKNYNTNLRLKLTXXI 120  
DB 61 RPIIKSGDEPDEVDILRLKLLGNRSLIAGRLPGRTANDVKNYNTNLRLKLTXXI 120  
QY 121 VPERKINNKGEISTKIEIKPQRRYFSSYMKNTNNVILDEEHCKEIISEKOTPA 180  
DB 121 VPERKINNKGEISTKIEIKPQRRYFSSYMKNTNNVILDEEHCKEIISEKOTPA 180  
QY 181 SMNNDPMTNINLENCNDIEDEEVINYEKTLTSLHEEISPPNIGSGMGOQIS 240  
DB 181 SMNNDPMTNINLENCNDIEDEEVINYEKTLTSLHEEISPPNIGSGMGOQIS 240  
QY 241 HENMGESFSLNLPMMQGVONDDPSAEIDMLNLLD 274  
DB 241 HENMGESFSLNLPMMQGVONDDPSAEIDMLNLLD 274

RESULT 2  
ID AAE01933 standard; Protein; 209 AA.

AC AAE01933;  
XX  
XX  
XX 31-JUL-2001 (first entry)  
XX  
XX Arabidopsis thaliana transcription factor homologue G2421.  
XX  
XX  
XX Transcription factor; trait modification; seed characteristic;  
XX structural characteristic; developmental characteristic; gene therapy;  
XX agricultural biotechnology.  
XX  
XX Arabidopsis thaliana.  
XX  
XX  
XX Key Location/Qualifiers  
XX FT Domain 9..110  
XX FT /Label= Conserved\_domain  
XX  
XX PN WO200135727-A1.  
XX  
XX PD 25-MAY-2001.  
XX  
XX PF 14-NOV-2000; 2000WO-US31457.  
XX  
XX PR 17-NOV-1999; 99US-0166228.  
XX  
XX PR 17-APR-2000; 2000US-0197899.  
XX  
XX PR 22-AUG-2000; 2000US-0227439.  
XX





DE Arabidopsis thaliana transcription factor homologue G2422.  
 XX  
 XX Transcription factor; trait modification; seed characteristic;  
 KM structural characteristic; developmental characteristic; gene therapy;  
 KM agricultural biotechnology.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 9..110  
 FT /label=Conserved\_domain  
 XX  
 PN WO200135727-A1.  
 XX  
 PD 25-MAY-2001.  
 XX  
 PF 14-NOV-2000; 2000WO-US31457.  
 XX  
 XX 17-NOV-1999; 99US-0166228.  
 PR 17-APR-2000; 2000US-0197899.  
 PR 22-AUG-2000; 2000US-0227439.  
 XX  
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (REUB/) REUBER L.  
 PA (CREE/) CREELMAN R.  
 PA (PILG/) PILGRIM M.  
 PA (RIEC/) RIECHMANN J L.  
 PA (JIANG/) JIANG C.  
 PA (YUGG/) YU G.  
 PA (PINE/) PINEDA O.  
 PA (HEAR/) HEARD J.  
 XX  
 PI Reuber L, Creelman R, Pilgrim M, Riechmann JL, Jiang C, Yu G;  
 PI Pineda O, Heard J;  
 DR WPI; 2001-335979/35.  
 DR N-PSDB; AAD05767.  
 XX  
 PT Nucleic acids encoding plant transcription factor polypeptides, useful  
 PT for altering the developmental and structural characteristics of  
 PT plants, e.g. corn, potato and cotton plants -  
 XX  
 PS Claim 4; Page 114; 133pp; English.  
 XX  
 CC The present sequence is homologue G2422 of Arabidopsis thaliana  
 CC Transcription factor G663. The transcription factors may be used to  
 CC modify traits associated with structural or developmental  
 CC characteristics of plants, e.g., corn, potato and cotton plants, when  
 CC their expression level is altered. Specifically, they are used for  
 CC modifying the nutritional content of plants or seeds, e.g., to modify  
 CC vitamin and mineral content, to modify the oil content of seeds, to  
 CC modify the insoluble sugar content of seeds, to modify phenyl lipid  
 CC content, to modify fatty acid and modify wax content. They may also be  
 CC used to alter seed characteristics such as shelf-life, size, stress  
 CC tolerance, seedling vigour, pest and pathogen resistance and germination  
 CC rate. They may also be used in gene therapy. Therefore manipulating  
 CC transcription factor levels in plants offers great potential in  
 CC agricultural biotechnology for modifying a plant's traits.  
 XX  
 SQ Sequence 246 AA;  
 XX  
 QY Query Match 11.7%; Score 32; DB 22; Length 246;  
 Db Best Local Similarity 100.0%; Pred. No. 3.9e-23;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 78 LRLHKLGNRWSLJAGRLPGRTANDVKNYNT 109  
 Db 75 LRLHKLGNRWSLJAGRLPGRTANDVKNYNT 106  
 RESULT 5  
 AAU93177  
 ID AAU93177 standard; Protein; 246 AA.

XX AAU93177;  
 AC 02-JUL-2002 (first entry)  
 DT  
 XX Arabidopsis transcription factor #215.  
 DE  
 XX Agriculture; metabolic chemical; environmental stress; drought;  
 KM microbial disease resistance; herbicide resistance; seed yield;  
 KM fruit yield; growth rate; leaf senescence; flower senescence.  
 KM plant; transcription factor; transgenic.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO200215675-A1.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 22-AUG-2001; 2001WO-US26189.  
 XX  
 XX 22-AUG-2000; 2000US-227439P.  
 PR 16-NOV-2000; 2000US-0713994.  
 PR 16-APR-2001; 2001US-0837944.  
 XX  
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (PILG/) PILGRIM M.  
 PA (CREE/) CREELMAN R.  
 PA (DUBE/) DUBELL A J.  
 PA (HEAR/) HEARD J.  
 PA (JIANG/) JIANG C.  
 PA (KEDD/) KEDDIE J.  
 PA (ADAM/) ADAM L.  
 PA (RATC/) RATCLIFF O.  
 PA (REUB/) REUBER J L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (YUGG/) YU G.  
 PA (PINE/) PINEDA O.  
 XX  
 PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;  
 PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
 DR WPI; 2002-292022/33.  
 DR N-PSDB; ABK65363.  
 XX  
 PT An isolated or recombinant polynucleotide used to produce a transgenic  
 PT plant -  
 XX  
 PS Claim 40; Page 880-881; 941pp; English.  
 XX  
 CC The invention relates to 1 of 232 isolated or recombinant polynucleotides  
 CC encoding an Arabidopsis thaliana transcription factor, their variants,  
 CC complements, fragments, or related polynucleotide with 31% to 95%  
 CC sequence identity, where the plant possesses an altered trait as compared  
 CC to a wild-type or reference plant, or the plant exhibits an altered  
 CC phenotype as compared to a wild-type or reference plant, or the plant  
 CC exhibits ecotypic expression or altered expression of one or more genes  
 CC associated with a plant trait as compared to a wild plant. Also included  
 CC are a transgenic plant comprising the polynucleotides, a computer  
 CC readable medium having stored sequence information, and identifying a  
 CC homologue sequence from a database comprising a plurality of known plant  
 CC sequences comprising inputting sequence information selected from one of  
 CC 464 fully defined sequences given in the specification. The isolated or  
 CC recombinant polynucleotide is used for producing a plant having a  
 CC modified trait, the method comprising selecting a polynucleotide that  
 CC encodes a polypeptide or an antisense nucleic acid, inserting the  
 CC polynucleotide or antisense nucleic acid into an expression vector,  
 CC introducing the vector into a plant or a cell of a plant to overexpress  
 CC the polypeptide or antisense nucleic acid, thereby producing a modified  
 CC plant, and selecting for a modified trait (e.g. increased  
 CC production of agriculturally useful proteins or metabolic chemicals,  
 CC pest tolerance, environmental stress response (e.g. drought), microbial  
 CC disease resistance, herbicide resistance, seed and fruit yield, growth  
 CC rate, leaf and flower senescence and many other traits listed in the

CC specification). The present sequence is one of the 232 proteins which are  
CC A. thaliana transcription factors.

SO Sequence 246 AA;

Query Match 11.7%; Score 32; DB 23; Length 246;  
Best Local Similarity 100.0%; Pred. No. 3.9e-23;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
Db 75 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 106

RESULT 6  
AAE01916

ID AAE01916 standard; Protein; 249 AA.

AC AAE01916;

DT 31-JUL-2001 (first entry)

DE Arabidopsis thaliana transcription factor G663.

KW Transcription factor; trait modification; seed characteristic;  
KW structural characteristic; developmental characteristic; gene therapy;  
KW agricultural biotechnology.

OS Arabidopsis thaliana.

Key Location/Qualifiers

FT Domain 9..111 /label=Conserved\_domain

PN WO200135727-A1.

PD 25-MAY-2001.

PF 14-NOV-2000; 2000MO-US31457.

PR 17-NOV-1999; 99US-0166228.

PR 17-APR-2000; 2000US-0197899.

PR 22-AUG-2000; 2000US-0227439.

PA (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (REUB/) REUBER L.

PA (CREE/) CREELMAN R.

PA (PILG/) PILGRIM M.

PA (RIEC/) RIECHMANN J L.

PA (JIANG/) JIANG C.

PA (YUGG/) YU G.

PA (PINE/) PINEDA O.

PA (HEAR/) HEARD J.

PI Reuber L, Creelman R, Pilgrim M, Riechmann JL, Jiang C, Yu G;

PI Pineda O, Heard J;

PS Claim 4; Page 70-71; 133pp; English.

CC The present sequence is Arabidopsis thaliana transcription factor G663.  
CC The transcription factors may be used to modify traits associated  
CC with structural or developmental characteristics of plants, e.g., corn,  
CC potato and cotton plants, when their expression level is altered.  
CC Specifically, they are used for modifying the nutritional content of  
CC plants or seeds, e.g., to modify vitamin and mineral content, to modify  
CC the oil content of seeds, to modify the insoluble sugar content of

CC seeds, to modify phenyl lipid content, to modify fatty acid and modify  
CC wax content. They may also be used to alter seed characteristics such  
CC as shelf-life, size, stress tolerance, seedling vigour, pest and  
CC pathogen resistance and germination rate. They may also be used in gene  
CC therapy. Therefore manipulating transcription factor levels in plants  
CC offers great potential in agricultural biotechnology for modifying a  
CC plant's traits.

SO Sequence 249 AA;

Query Match 11.7%; Score 32; DB 22; Length 249;  
Best Local Similarity 100.0%; Pred. No. 3.9e-23;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
Db 75 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 106

RESULT 7  
AAB68356

ID AAB68356 standard; Protein; 249 AA.

AC AAB68356;

DT 09-JUL-2001 (first entry)

DE Amino acid sequence of MYB transcription factor AtMYB90.

KW Stress tolerance; myloblastosis transcription factor; heat tolerance;  
KW MYB transcription factor; AtMYB60; AtMYB764; AtMYB75; transgenic plant;  
KW phenylpropanoid biosynthesis.

OS Arabidopsis sp.

PN WO200132002-A1.

PD 10-MAY-2001.

PF 06-NOV-2000; 2000MO-US30503.

PR 05-NOV-1999; 99US-0163579.

PR 23-OCT-2000; 2000US-0693855.

PA (BADI ) BASF CORP.

PA Tonelli C;

DR WPI; 2001-316365/73.

DR N-PSDB; AAF65191.

XX Nucleic acid molecules encoding plant stress tolerance-related

PT myloblastosis transcription factors for increasing stress resistance of

PS a crop in a field -

PS Claim 13; Fig 4; 92pp; English.

CC The present sequence represents a plant stress tolerance-related  
CC myloblastosis (MYB) transcription factor. The specification describes  
CC AtMYB60, AtMYB764, AtMYB75 and AtMYB90. MYB polypeptides are useful for  
CC enhancing a plant's tolerance or sensitivity to stress e.g. salt,  
CC drought, cold and heat tolerance. It is also useful for producing a  
CC transgenic plant with enhanced stress tolerance or stress sensitivity.  
CC MYB polypeptides are also useful for increasing the production of  
CC products of the phenylpropanoid biosynthesis pathway, e.g. stilbenes,  
CC flavonoids, lignins, salicylic acid, anthocyanins, and phenolic  
CC derivatives.

SO Sequence 249 AA;

Query Match 11.7%; Score 32; DB 22; Length 249;  
Best Local Similarity 100.0%; Pred. No. 3.9e-23;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LRLHKLGNRMSLIAGRLPGRANDVKNYWT 109  
 |||||||  
 DB 75 LRLHKLGNRMSLIAGRLPGRANDVKNYWT 106

RESULT 8  
 AAU92969 standard; Protein: 249 AA.  
 AC AAU92969;  
 XX  
 DT 02-JUL-2002 (first entry)  
 XX  
 DE Arabidopsis transcription factor #7.  
 XX  
 KM Agriculture; metabolic chemical; environmental stress; drought;  
 KM microbial disease resistance; herbicide resistance; seed yield;  
 KM fruit yield; growth rate; leaf senescence; flower senescence.  
 KM plant; transcription factor; transgenic.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO200215675-A1.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 22-AUG-2001; 2001WO-US26189.  
 XX  
 PR 22-AUG-2000; 2000US-227439P.  
 PR 16-NOV-2000; 2000US-0713994.  
 PR 16-APR-2001; 2001US-0837944.  
 XX  
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (PILG/) PILGRIM M.  
 PA (CREE/) CREELMAN R.  
 PA (DUBE/) DUBBEL A J.  
 PA (HEAR/) HEARD J.  
 PA (JIANG/) JIANG C.  
 PA (KEDD/) KEDDIE J.  
 PA (ADAM/) ADAM L.  
 PA (RATC/) RATCLIFF O.  
 PA (REUB/) REUBER J L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (YUGG/) YU G.  
 PA (PINE/) PINEDA O.  
 XX  
 PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J,  
 PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
 XX  
 DR WPI: 2002-297022/33.  
 DR N-PSDB; ABK65155.  
 XX  
 PT An isolated or recombinant polynucleotide used to produce a transgenic  
 PT plant -  
 XX  
 PS Claim 40; Page 97-98; 941pp; English.  
 XX  
 CC The invention relates to 1 of 232 isolated or recombinant polynucleotides  
 CC encoding an Arabidopsis thaliana transcription factor, their variants,  
 CC complements, fragments, or related polynucleotide with 31% to 95%  
 CC sequence identity, where the plant possesses an altered trait as compared  
 CC to a wild-type or reference plant, or the plant exhibits an altered  
 CC phenotype as compared to a wild-type or reference plant, or the plant  
 CC exhibits ectopic expression or altered expression of one or more genes  
 CC associated with a plant trait as compared to a wild plant. Also included  
 CC are a transgenic plant comprising the polynucleotide, a computer  
 CC readable medium having stored sequence information, and identifying a  
 CC homologue sequence from a database comprising a plurality of known plant  
 CC sequences comprising inputting sequence information selected from one of  
 CC 464 fully defined sequences given in the specification. The isolated or  
 CC recombinant polynucleotide is used for producing a plant having a  
 CC modified trait, the method comprising selecting a polynucleotide that

CC encodes a polypeptide or an antisense nucleic acid, inserting the  
 CC polynucleotide or antisense nucleic acid into an expression vector,  
 CC introducing the vector into a plant or a cell of a plant to overexpress  
 CC the polypeptide or antisense nucleic acid, thereby producing a modified  
 CC plant, and selecting for a modified trait (e.g. increased  
 CC production of agriculturally useful proteins or metabolic chemicals,  
 CC pest tolerance, environmental stress response (e.g. drought), microbial  
 CC disease resistance, herbicide resistance, seed and fruit yield, growth  
 CC rate, leaf and flower senescence and many other traits listed in the  
 CC specification). The present sequence is one of the 232 proteins which are  
 CC A. thaliana transcription factors.  
 CC  
 XX  
 SQ Sequence 249 AA;  
 XX  
 QY 78 LRLHKLGNRMSLIAGRLPGRANDVKNYWT 109  
 |||||||  
 DB 75 LRLHKLGNRMSLIAGRLPGRANDVKNYWT 106

Query Match 11.7%; Score 32; DB 23; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-23;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 9  
 AAU75735 standard; Protein: 249 AA.  
 AC AAU75735;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE A. thaliana Production of anthocyanin pigment 2 (PAP2) protein.  
 XX  
 KM PAP2; production of anthocyanin pigment; MYB-like transcription factor;  
 KM transgenic; plant; phenylpropanoid; stress; light stress; water stress;  
 KM pH stress; temperature stress; heavy metal stress; pathogen attack;  
 KM infection; wounding; nutrient deficiency; herbivory; plant colouration;  
 KM stress sensitisation; chromosome 1.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_difference 130  
 FT /note= "Encoded by CCT"  
 FT misc\_difference 161  
 FT /note= "Encoded by GAA"  
 FT misc\_difference 190  
 FT /note= "Encoded by GAT"  
 FT misc\_difference 215  
 FT /note= "Encoded by GAT"  
 FT misc\_difference 244  
 FT /note= "Encoded by GAG"  
 XX  
 PN WO200200902-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 21-JUN-2001; 2001WO-US19734.  
 XX  
 PR 23-JUN-2000; 2000US-0603244.  
 PR 05-JUL-2000; 2000US-0610185.  
 XX  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 PA (ROBE-) ROBERTS NOBLE FOUND INC SAMMUEL.  
 PI Borevitz J, Xia Y, Lamb CJ, Dixon RA;  
 DR WPI: 2002-16443/21.  
 DR N-PSDB; ABK14222.  
 XX  
 PT New isolated PAP1 or PAP2 gene, useful for increasing pigmentation in  
 PT plants, as reporter genes for analysing expression pattern of promoter  
 PT of interest, and to increase flux through phenylpropanoid pathway -

XX Claim 9; Page 28; 29pp; English.  
PS  
XX  
CC This invention relates to the nucleotide and protein sequences of novel  
CC Arabidopsis thaliana production of anthocyanin pigment (PAP) genes PAP1  
CC or PAP2. The genes for both PAP1 and PAP2 have been mapped to chromosome  
CC 1. PAP1 and PAP2 are MYB-like transcription factors that regulate the  
CC production of anthocyanin pigment. The nucleotide sequence of the  
CC invention is useful for screening for PAP gene expression. It is also  
CC useful for detecting promoter activity within a plant cell. The  
CC nucleotide sequence is also useful for screening for plant cell (e.g.,  
CC A. thaliana) transformation. The transformed plant cells are then used  
CC to produce transgenic plants with tissue-specific altered colour  
CC expression. The nucleotide sequence is also useful for enhancing  
CC accumulation of phenylpropanoid product in a plant. A transgenic  
CC plant is useful for detecting a stress condition such as light  
CC stress, water stress, pH stress, temperature stress, heavy metal  
CC stress, pathogen attack or infection, wounding, nutrient deficiency,  
CC herbivory, or abnormal hormone levels in a plant. PAP1 or PAP2  
CC overexpression can also be useful in developing or enhancing desirable  
CC colouration in ornamental plants for commercial sales, and for  
CC sensitizing plants to stress. Introduction of PAP1 or PAP2 sensitises  
CC plants to respond more quickly to a stress condition in ways that are  
CC physiologically significant, as well as in ways of ornamental  
CC significance. The present sequence represents the A. thaliana PAP2  
CC protein sequence of the invention.  
XX  
SQ Sequence 249 AA;  
Query Match 11.7%; Score 32; DB 23; Length 249;  
Best Local Similarity 100.0%; Pred. No. 3.9e-23;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 78 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
DB 75 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 106  
RESULT 10  
AAB30046  
ID AAB30046 standard; Protein; 249 AA.  
XX  
AC AAE30046;  
XX  
DT 24-FEB-2003 (first entry)  
XX  
DE Arabidopsis thaliana G663 transcription factor protein.  
XX  
KM Transcription factor; metabolite pathway; terpenoid; limonene synthase;  
KM alkaloid pathway gene; taxadiene synthase; biological pathway; freezing;  
KM abiotic stress; cold; drought; heat; nutrient deficiency; biotic stress;  
KM infection; developmental pathway; flowering; root development; TDS; US;  
KM transgenic; transgenic plant.  
XX  
OS Arabidopsis thaliana.  
XX  
PN W0200274917-A2.  
XX  
PD 26-SEP-2002.  
XX  
PF 15-MAR-2002; 2002MO-US07999.  
XX  
PR 16-MAR-2001; 2001US-0810836.  
XX  
PA (MEND-) MENDEL BIOTECHNOLOGY INC.  
XX  
PI Brown P;  
XX  
DR WPI; 2003-018799/01.  
XX  
DR N-PSDB; AAD47497.  
XX  
PT Determining whether one of several test transcription factor (TF)  
PT polynucleotides encodes pathway TF by determining expression from

PT pathway gene promoter linked to reporter gene in a cell in presence of  
PT test polynucleotides  
XX  
PS Disclosure; Page 153; 221pp; English.  
XX  
XX The present invention relates to a high-throughput method for identifying  
CC a polynucleotide which encodes a transcription factor for controlling the  
CC expression of one or more genes in a pathway. The method is useful for  
CC determining whether a member of a pool of test transcription factor or  
CC polynucleotides encodes a biosynthetic pathway transcription factor or  
CC a primary metabolite pathway gene or a secondary metabolite pathway gene  
CC such as a terpenoid or alkaloid pathway gene transcription factor. It  
CC is also useful for determining whether a member of a pool of test encodes  
CC a mentha sp. or Taxus sp. terpenoid pathway gene where the terpenoid  
CC pathway gene encodes limonene synthase (LS) or taxadiene synthase (TDS).  
CC It is also useful for identifying one or more transcription factors that  
CC activate one or more genes of a biological pathway of a plant, fungi or  
CC animal cell, where the biological pathway can be a biochemical pathway  
CC such as biosynthetic pathways for soluble and insoluble carbohydrates,  
CC for amino acids and a response pathway to abiotic stress such as cold,  
CC freezing, drought, heat, nutrient deficiency, pH or biotic stress such  
CC as fungal, viral or bacterial infection, developmental pathway such as  
CC flowering, root development, a response pathway to environmental cues  
CC such as light intensity and light quality, circadian rhythm. Sequences  
CC of the invention are used to generate transgenic plants. The present  
CC sequence is Arabidopsis thaliana G663 transcription factor protein.  
XX  
SQ Sequence 249 AA;  
Query Match 11.7%; Score 32; DB 24; Length 249;  
Best Local Similarity 100.0%; Pred. No. 3.9e-23;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 78 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
DB 75 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 106  
RESULT 11  
AAB68355  
ID AAB68355 standard; Protein; 211 AA.  
XX  
AC AAB68355;  
XX  
DT 09-JUL-2001 (first entry)  
XX  
DE Amino acid sequence of MYB transcription factor AtMYB75.  
XX  
KM Stress tolerance; myloblastosis transcription factor; heat tolerance;  
KM MYB transcription factor; AtMYB60; AtMYB764; AtMYB75; transgenic plant;  
KM phenylpropanoid biosynthesis.  
XX  
OS Arabidopsis sp.  
XX  
FH Key Location/Qualifiers  
FT MISC-difference 72 /note= "asp encoded by TTTCGAT"  
FT MISC-difference 90 /note= "Gly encoded by TTTGGA"  
FT MISC-difference 173 /note= "Asn encoded by TAAT"  
FT MISC-difference 180 /note= "His encoded by ATC"  
FT MISC-difference 181 /note= "Tyr encoded by ATATAC"  
XX  
PN W0200132002-A1.  
XX  
PD 10-MAY-2001.  
XX  
PF 06-NOV-2000; 2000MO-US30503.  
XX  
PR 05-NOV-1999; 99US-0163579.

PR 23-OCT-2000; 2000US-0693855.  
 XX (BADI ) BASF CORP.  
 PA  
 XX  
 PI Tonelli C;  
 XX  
 DR WPI; 2001-316365/33.  
 DR N-PSDB; AAF85190.  
 XX  
 PT Nucleic acid molecules encoding plant stress tolerance-related  
 PT myloblastosis transcription factors for increasing stress resistance of  
 a crop in a field -  
 PS Claim 13; Fig 3B; 92pp; English.  
 XX  
 CC The present sequence represents a plant stress tolerance-related  
 CC myloblastosis (MYB) transcription factor. The specification describes  
 CC AWMYB80, AWMYB64, AWMYB75 and AWMYB90. MYB polypeptides are useful for  
 CC enhancing a plant's tolerance or sensitivity to stress e.g. salt,  
 CC drought, cold and heat tolerance. It is also useful for producing a  
 CC transgenic plant with enhanced stress tolerance or stress sensitivity.  
 CC MYB polypeptides are also useful for increasing the production of  
 CC products of the phenylpropanoid biosynthesis pathway, e.g. stilbenes,  
 CC flavonoids, lignins, salicylic acid, anthocyanins, and phenolic  
 CC derivatives.  
 CC  
 SQ Sequence 211 AA;  
 Query Match 9.9%; Score 27; DB 22; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 3e-18;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 83 LAGNRWSLIAGRLPGRTANDVKNYNT 109  
 Db 80 LAGNRWSLIAGRLPGRTANDVKNYNT 106  
 RESULT 12  
 AAU93155  
 ID AAU93155 standard; Protein; 212 AA.  
 XX  
 AC AAU93155;  
 XX  
 DT 02-JUL-2002 (first entry)  
 XX  
 DE Arabidopsis transcription factor #193.  
 XX  
 KW Agriculture; metabolic chemical; environmental stress; drought;  
 KW microbial disease resistance; herbicide resistance; seed yield;  
 KW fruit yield; growth rate; leaf senescence; flower senescence.  
 KM plant; transcription factor; transgenic.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO200215675-A1.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 22-AUG-2001; 2001WO-US26189.  
 XX  
 PR 22-AUG-2000; 2000US-227439P.  
 PR 16-NOV-2000; 2000US-0713994.  
 PR 16-APR-2001; 2001US-0837944.  
 XX  
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (PIG/) PILGRIM M.  
 PA (CREE/) CREELMAN R.  
 PA (DUBE/) DUBELL A J.  
 PA (HEAR/) HEARD J.  
 PA (JIANG/) JIANG C.  
 PA (KEDD/) KEDDIE J.  
 PA (ADAM/) ADAM L.  
 PA (RATC/) RATCLIFF O.

PA (REUB/) REUBER J L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (YUGG/) YU G.  
 PA (PINE/) PINEDA O.  
 XX  
 PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;  
 PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
 XX  
 DR WPI; 2002-292022/33.  
 DR N-PSDB; ABK65341.  
 XX  
 PT An isolated or recombinant polynucleotide used to produce a transgenic  
 PT plant -  
 PS Claim 40; Page 794-795; 941pp; English.  
 XX  
 CC The invention relates to 1 of 232 isolated or recombinant polynucleotides  
 CC encoding an Arabidopsis thaliana transcription factor, their variants,  
 CC complements, fragments, or related polynucleotide with 31% to 95%  
 CC sequence identity, where the plant possesses an altered trait as compared  
 CC to a wild-type or reference plant, or the plant exhibits an altered  
 CC phenotype as compared to a wild-type or reference plant, or the plant  
 CC exhibits ectopic expression or altered expression of one or more genes  
 CC associated with a plant trait as compared to a wild plant. Also included  
 CC are a transgenic plant comprising the polynucleotides, a computer  
 CC readable medium having stored sequence information, and identifying a  
 CC homologue sequence from a database comprising a plurality of known plant  
 CC sequences comprising inputting sequence information selected from one of  
 CC 464 fully defined sequences given in the specification. The isolated or  
 CC recombinant polynucleotide is used for producing a plant having a  
 CC modified trait, the method comprising selecting a polynucleotide that  
 CC encodes a polypeptide or an antisense nucleic acid, inserting the  
 CC polynucleotide or antisense nucleic acid into an expression vector,  
 CC introducing the vector into a plant or a cell of a plant to overexpress  
 CC the polypeptide or antisense nucleic acid, thereby producing a modified  
 CC plant, and selecting for a modified trait (e.g. increased  
 CC production of agriculturally useful proteins or metabolic chemicals,  
 CC pest tolerance, environmentally stress response (e.g. drought), microbial  
 CC disease resistance, herbicide resistance, seed and fruit yield, growth  
 CC rate, leaf and flower senescence and many other traits listed in the  
 CC specification). The present sequence is one of the 232 proteins which are  
 CC A. thaliana transcription factors.  
 CC  
 SQ Sequence 212 AA;  
 Query Match 9.9%; Score 27; DB 23; Length 212;  
 Best Local Similarity 100.0%; Pred. No. 3e-18;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 83 LAGNRWSLIAGRLPGRTANDVKNYNT 109  
 Db 80 LAGNRWSLIAGRLPGRTANDVKNYNT 106  
 RESULT 13  
 AAU75734  
 ID AAU75734 standard; Protein; 248 AA.  
 XX  
 AC AAU75734;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE A. thaliana Production of anthocyanin pigment 1 (PAP1) protein.  
 XX  
 KW PAP1; production of anthocyanin pigment; MYB-like transcription factor;  
 KW transgenic; plant; phenylpropanoid; stress; light stress; water stress;  
 KW pH stress; temperature stress; heavy metal stress; pathogen attack;  
 KW infection; wounding; nutrient deficiency; herbivory; plant colouration;  
 KW stress sensitisation; chromosome 1.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO200200902-A2.



XX Wood M, McGrath A, Shenk MA, Glenn M;  
 XX WPI; 2000-579369/54.

XX New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide

XX Claim 8; Page 726; 747pp; English.

XX The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such  
 CC transcription factor. The transcription factor may be used to produce a  
 CC plant having modified gene expression such as a woody plant e.g. a  
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or manogany species or  
 CC to modify the activity of a polypeptide in a plant. The transcription  
 CC factors of the present invention are members from the following families  
 CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic  
 CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain  
 CC zipper, LIM domain, AP2 and ERFs, zinc finger domains of type 2  
 CC Cys2His2, CCAAT box elements and MYB.

XX Sequence 48 AA;

Query Match 6.2%; Score 17; DB 21; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-09;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 LLGNRMSLIAGRLLPGRT 99  
 |||||  
 Db 22 LLGNRMSLIAGRLLPGRT 38

Search completed: January 29, 2004, 20:19:33  
 Job time : 57 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 20:18:28 ; Search time 21 Seconds  
(without alignments)  
552.056 Million cell updates/sec

Title: US-10-033-190-2

Perfect score: 274

Sequence: 1 MNSTMSLSIGVRKGSWTDEE.....QQGVQNDPSAIDLNWLLD 274

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	11.7	249	4	US-09-610-185C-4
2	27	9.9	248	4	US-09-610-185C-2
3	8	2.9	1247	1	US-08-472-934-10
4	8	2.9	1247	2	US-08-323-460A-10
5	8	2.9	1247	2	US-08-461-146C-10
6	8	2.9	1247	3	US-08-461-145C-10
7	8	2.9	1597	4	US-09-423-890-13
8	8	2.9	1597	4	US-08-628-829-14
9	7	2.6	60	3	US-09-058-489-4
10	7	2.6	183	4	US-09-252-991A-20768
11	7	2.6	218	4	US-09-252-991A-26955
12	7	2.6	278	2	US-08-722-626B-2
13	7	2.6	289	4	US-09-328-352-6229
14	7	2.6	301	4	US-09-495-406-15
15	7	2.6	324	2	US-08-671-320-11
16	7	2.6	324	2	US-08-671-320-13
17	7	2.6	324	2	US-08-671-320-11
18	7	2.6	324	2	US-08-671-320-13
19	7	2.6	324	2	US-08-671-320-11
20	7	2.6	324	2	US-08-671-320-13
21	7	2.6	419	4	US-09-252-991A-17834
22	7	2.6	445	4	US-09-252-991A-23505
23	7	2.6	464	4	US-09-252-991A-26212
24	7	2.6	464	4	US-09-252-991A-28286
25	7	2.6	553	3	US-08-997-251-2
26	7	2.6	553	3	US-08-997-251-4
27	7	2.6	553	4	US-09-252-991A-17984

28	7	2.6	580	4	US-09-252-991A-30180	Sequence 30180, A
29	7	2.6	708	3	US-09-413-814-18	Sequence 18, App1
30	7	2.6	830	4	US-09-562-737-32	Sequence 32, App1
31	7	2.6	844	4	US-09-564-805-227	Sequence 227, App
32	7	2.6	1247	4	US-09-252-991A-32960	Sequence 32960, A
33	6	2.2	10	4	US-09-081-975-17	Sequence 17, App1
34	6	2.2	15	6	5439796-2	Patent No. 5439796
35	6	2.2	15	6	5439796-3	Patent No. 5439796
36	6	2.2	20	3	US-09-030-267-1	Sequence 1, App1
37	6	2.2	20	4	US-09-081-975-15	Sequence 15, App1
38	6	2.2	20	4	US-09-081-975-15	Sequence 15, App1
39	6	2.2	29	4	US-09-227-357-404	Sequence 404, App
40	6	2.2	50	3	US-09-156-316-5	Sequence 5, App1
41	6	2.2	51	1	US-08-519-103-19	Sequence 19, App1
42	6	2.2	51	3	US-09-018-635-19	Sequence 19, App1
43	6	2.2	51	4	US-09-312-962-19	Sequence 19, App1
44	6	2.2	52	1	US-08-519-103-13	Sequence 13, App1
45	6	2.2	52	3	US-09-018-635-13	Sequence 13, App1
46	6	2.2	52	4	US-09-912-962-13	Sequence 13, App1
47	6	2.2	59	1	US-08-470-179-1	Sequence 1, App1
48	6	2.2	59	1	US-08-470-179-2	Sequence 2, App1
49	6	2.2	73	4	US-08-469-260A-510	Sequence 510, App
50	6	2.2	73	4	US-08-488-446-510	Sequence 510, App
51	6	2.2	73	4	US-08-467-344A-510	Sequence 510, App
52	6	2.2	95	4	US-09-309-572-8	Sequence 8, App1
53	6	2.2	102	3	US-08-478-097A-18	Sequence 18, App1
54	6	2.2	102	4	US-09-496-398-18	Sequence 18, App1
55	6	2.2	103	4	US-09-198-452A-1230	Sequence 1230, App
56	6	2.2	119	4	US-09-252-991A-22545	Sequence 22545, A
57	6	2.2	133	4	US-09-252-991A-16979	Sequence 16979, A
58	6	2.2	134	4	US-09-252-991A-21251	Sequence 21251, A
59	6	2.2	134	4	US-09-732-210-107	Sequence 107, App
60	6	2.2	137	4	US-09-732-210-572	Sequence 572, App
61	6	2.2	141	4	US-09-732-210-376	Sequence 376, App
62	6	2.2	142	4	US-09-732-210-641	Sequence 641, App
63	6	2.2	143	4	US-09-252-991A-30385	Sequence 30385, A
64	6	2.2	145	4	US-09-252-991A-20032	Sequence 20032, A
65	6	2.2	148	4	US-09-252-991A-22571	Sequence 22571, A
66	6	2.2	152	4	US-09-252-991A-32584	Sequence 32584, A
67	6	2.2	155	4	US-09-252-991A-17230	Sequence 17230, A
68	6	2.2	158	3	US-09-042-071-20	Sequence 20, App1
69	6	2.2	159	4	US-09-621-377B-4	Sequence 4, App1
70	6	2.2	162	4	US-09-252-991A-21465	Sequence 21465, A
71	6	2.2	164	4	US-09-252-991A-21892	Sequence 21892, A
72	6	2.2	166	4	US-09-252-991A-20684	Sequence 20684, A
73	6	2.2	169	1	US-08-534-975-2	Sequence 2, App1
74	6	2.2	169	2	US-08-534-975-2	Sequence 2, App1
75	6	2.2	169	3	US-09-129-855A-2	Sequence 2, App1
76	6	2.2	169	3	US-09-247-154-2	Sequence 2, App1
77	6	2.2	169	3	US-09-480-718-2	Sequence 2, App1
78	6	2.2	169	4	US-09-610-833-2	Sequence 2, App1
79	6	2.2	169	4	US-09-129-855A-2	Sequence 2, App1
80	6	2.2	191	4	US-09-252-991A-18479	Sequence 18479, A
81	6	2.2	196	4	US-09-227-357-393	Sequence 393, App
82	6	2.2	196	4	US-09-198-452A-1204	Sequence 1204, App
83	6	2.2	197	2	US-08-923-867-2	Sequence 2, App1
84	6	2.2	197	3	US-08-928-100-4	Sequence 4, App1
85	6	2.2	197	3	US-09-183-134-2	Sequence 2, App1
86	6	2.2	197	4	US-09-492-581-4	Sequence 4, App1
87	6	2.2	197	4	US-09-425-666-4	Sequence 4, App1
88	6	2.2	206	4	US-09-166-350-21	Sequence 21, App1
89	6	2.2	208	3	US-08-896-933-30	Sequence 30, App1
90	6	2.2	208	4	US-09-314-235-30	Sequence 30, App1
91	6	2.2	209	4	US-09-252-991A-27925	Sequence 27925, A
92	6	2.2	211	4	US-09-582-379-5	Sequence 5, App1
93	6	2.2	212	4	US-09-252-991A-23200	Sequence 23200, A
94	6	2.2	220	4	US-09-252-991A-21649	Sequence 21649, A
95	6	2.2	224	4	US-09-328-352-6616	Sequence 6616, App
96	6	2.2	226	4	US-09-327-357-401	Sequence 401, App
97	6	2.2	242	4	US-09-252-991A-19868	Sequence 19868, A
98	6	2.2	248	3	US-08-935-263-14	Sequence 14, App1
99	6	2.2	248	4	US-09-594-185-14	Sequence 14, App1
100	6	2.2	248	4	US-09-252-991A-28814	Sequence 28814, A

101	6	2.2	249	1	US-08-450-257-57	Sequence 57, Appl	174	6	2.2	376	4	US-09-328-352-7587	Sequence 7587, Ap
102	6	2.2	249	1	US-08-450-246-57	Sequence 57, Appl	175	6	2.2	381	4	US-09-252-991A-17106	Sequence 17106, A
103	6	2.2	249	1	US-08-450-098-57	Sequence 57, Appl	176	6	2.2	385	3	US-09-413-814-7	Sequence 7, Appl
104	6	2.2	249	1	US-08-451-233-57	Sequence 57, Appl	177	6	2.2	386	4	US-09-252-991A-25124	Sequence 25124, A
105	6	2.2	249	1	US-08-450-236-57	Sequence 57, Appl	178	6	2.2	389	4	US-09-107-532A-6185	Sequence 6185, Ap
106	6	2.2	249	1	US-08-235-403-57	Sequence 17, Appl	179	6	2.2	392	3	US-09-416-050A-2	Sequence 2, Appl
107	6	2.2	250	3	US-08-968-563-17	Sequence 17, Appl	180	6	2.2	392	3	US-09-664-800-2	Sequence 2, Appl
108	6	2.2	250	3	US-08-969-683A-17	Sequence 17, Appl	181	6	2.2	392	3	US-09-665-309-2	Sequence 2, Appl
109	6	2.2	250	4	US-09-297-928-13	Sequence 13, Appl	182	6	2.2	392	3	US-09-661-569-2	Sequence 2, Appl
110	6	2.2	250	4	US-09-134-001C-5439	Sequence 5439, Ap	183	6	2.2	396	4	US-09-252-991A-28645	Sequence 28645, A
111	6	2.2	250	4	US-09-641-652-56	Sequence 56, Appl	184	6	2.2	398	4	US-09-134-001C-5369	Sequence 5369, Ap
112	6	2.2	259	4	US-09-522-714-18	Sequence 18, Appl	185	6	2.2	399	1	US-08-480-528A-2	Sequence 2, Appl
113	6	2.2	262	4	US-09-186-276B-35	Sequence 35, Appl	186	6	2.2	399	1	US-08-479-666-2	Sequence 2, Appl
114	6	2.2	262	4	US-08-842-445-35	Sequence 35, Appl	187	6	2.2	399	1	US-08-459-346-9	Sequence 9, Appl
115	6	2.2	262	4	US-09-186-188B-35	Sequence 35, Appl	188	6	2.2	399	2	US-08-901-200A-2	Sequence 2, Appl
116	6	2.2	263	2	US-08-972-008-2	Sequence 2, Appl	189	6	2.2	399	3	US-08-889-419-9	Sequence 9, Appl
117	6	2.2	263	4	US-09-141-027-2	Sequence 2, Appl	190	6	2.2	399	3	US-09-219-391-2	Sequence 2, Appl
118	6	2.2	263	4	US-09-267-409-2	Sequence 2, Appl	191	6	2.2	399	3	US-08-808-346-4	Sequence 4, Appl
119	6	2.2	263	4	US-09-617-80A-2	Sequence 2, Appl	192	6	2.2	399	4	US-08-402-542-9	Sequence 9, Appl
120	6	2.2	266	4	US-09-252-991A-32392	Sequence 32392, A	193	6	2.2	399	4	US-08-643-321-26	Sequence 26, Appl
121	6	2.2	270	4	US-09-328-352-6871	Sequence 6871, Ap	194	6	2.2	399	4	US-09-328-352-7334	Sequence 7334, Ap
122	6	2.2	271	3	US-08-968-563-33	Sequence 33, Appl	195	6	2.2	399	5	PCT-US93-07189-9	Sequence 9, Appl
123	6	2.2	271	3	US-08-969-683A-33	Sequence 33, Appl	196	6	2.2	399	5	PCT-US93-10520-2	Sequence 2, Appl
124	6	2.2	271	4	US-09-297-928-14	Sequence 14, Appl	197	6	2.2	402	4	US-09-252-991A-28501	Sequence 28501, A
125	6	2.2	273	1	US-08-221-750A-9	Sequence 9, Appl	198	6	2.2	404	4	US-09-252-991A-25830	Sequence 25830, A
126	6	2.2	273	4	US-09-252-991A-16693	Sequence 16693, A	199	6	2.2	410	2	US-08-472-666-2	Sequence 2, Appl
127	6	2.2	274	4	US-09-252-991A-23361	Sequence 23361, A	200	6	2.2	410	4	US-09-101-068-4	Sequence 4, Appl
128	6	2.2	276	2	US-07-857-224B-72	Sequence 72, Appl	201	6	2.2	410	5	PCT-US96-07615-2	Sequence 2, Appl
129	6	2.2	277	4	US-09-186-276B-34	Sequence 34, Appl	202	6	2.2	415	4	US-09-252-991A-30711	Sequence 30711, A
130	6	2.2	277	4	US-08-842-445-34	Sequence 34, Appl	203	6	2.2	418	3	US-09-030-267-5	Sequence 5, Appl
131	6	2.2	277	4	US-09-186-188B-34	Sequence 34, Appl	204	6	2.2	422	4	US-09-252-991A-27513	Sequence 27513, A
132	6	2.2	277	4	US-09-252-991A-18435	Sequence 18435, A	205	6	2.2	424	4	US-09-252-991A-17882	Sequence 17882, A
133	6	2.2	284	4	US-09-914-359-43	Sequence 43, Appl	206	6	2.2	429	4	US-09-252-991A-23556	Sequence 23556, A
134	6	2.2	286	4	US-09-134-001C-3059	Sequence 3059, Ap	207	6	2.2	431	4	US-09-252-991A-20267	Sequence 20267, A
135	6	2.2	294	4	US-09-252-991A-32345	Sequence 32345, A	208	6	2.2	433	4	US-09-352-991A-21838	Sequence 21838, A
136	6	2.2	299	4	US-09-134-001C-4414	Sequence 4414, Ap	209	6	2.2	434	4	US-09-252-991A-32405	Sequence 32405, A
137	6	2.2	302	4	US-09-252-991A-22508	Sequence 22508, A	210	6	2.2	454	4	US-09-252-991A-32928	Sequence 32928, A
138	6	2.2	304	1	US-07-851-976B-8	Sequence 8, Appl	211	6	2.2	456	4	US-09-252-991A-17335	Sequence 17335, A
139	6	2.2	304	1	US-08-291-609-8	Sequence 8, Appl	212	6	2.2	457	4	US-09-252-991A-29798	Sequence 29798, A
140	6	2.2	304	1	US-08-401-136-8	Sequence 8, Appl	213	6	2.2	462	3	US-08-801-344-9	Sequence 9, Appl
141	6	2.2	306	3	US-08-850-554-8	Sequence 8, Appl	214	6	2.2	462	4	US-09-498-599-9	Sequence 9, Appl
142	6	2.2	306	4	US-09-252-991A-25699	Sequence 25699, A	215	6	2.2	462	4	US-09-252-991A-31372	Sequence 31372, A
143	6	2.2	307	4	US-09-252-991A-22508	Sequence 22508, A	216	6	2.2	466	4	US-09-252-991A-16891	Sequence 16891, A
144	6	2.2	308	4	US-09-252-991A-20925	Sequence 20925, A	217	6	2.2	467	4	US-09-328-352-7510	Sequence 7510, Ap
145	6	2.2	312	1	US-08-094-128A-27	Sequence 27, Appl	218	6	2.2	483	4	US-09-252-991A-19015	Sequence 19015, A
146	6	2.2	312	1	US-08-455-674-27	Sequence 27, Appl	219	6	2.2	486	4	US-09-252-991A-31407	Sequence 31407, A
147	6	2.2	312	1	US-08-455-992-27	Sequence 27, Appl	220	6	2.2	506	4	US-09-352-991A-27119	Sequence 27119, A
148	6	2.2	312	1	US-08-455-972-27	Sequence 27, Appl	221	6	2.2	512	4	US-08-999-689A-7	Sequence 7, Appl
149	6	2.2	312	1	US-08-252-991A-25338	Sequence 25338, A	222	6	2.2	515	4	US-09-252-991A-33046	Sequence 33046, A
150	6	2.2	312	5	PCT-US92-00652-27	Sequence 27, Appl	223	6	2.2	522	4	US-09-328-352-6840	Sequence 6840, Ap
151	6	2.2	313	3	US-09-134-758-4	Sequence 4, Appl	224	6	2.2	532	4	US-09-117-853-2	Sequence 2, Appl
152	6	2.2	313	4	US-09-996-243-414	Sequence 414, Ap	225	6	2.2	532	4	US-09-911-154-2	Sequence 2, Appl
153	6	2.2	313	4	US-09-709-677-4	Sequence 4, Appl	226	6	2.2	536	2	US-08-890-094-2	Sequence 2, Appl
154	6	2.2	315	4	US-09-227-357-155	Sequence 155, Ap	227	6	2.2	538	4	US-09-252-991A-19476	Sequence 19476, A
155	6	2.2	315	4	US-09-134-001C-4141	Sequence 4141, Ap	228	6	2.2	539	2	US-09-304-711-20	Sequence 20, Appl
156	6	2.2	320	4	US-09-252-991A-31701	Sequence 31701, A	229	6	2.2	539	4	US-09-304-711-20	Sequence 20, Appl
157	6	2.2	323	4	US-09-252-991A-16975	Sequence 16975, A	230	6	2.2	539	4	US-09-173-281-20	Sequence 20, Appl
158	6	2.2	323	4	US-09-252-991A-26948	Sequence 26948, A	231	6	2.2	548	2	US-08-890-094-18	Sequence 18, Appl
159	6	2.2	324	4	US-09-252-991A-26948	Sequence 26948, A	232	6	2.2	548	2	US-08-890-094-18	Sequence 18, Appl
160	6	2.2	334	4	US-09-252-991A-20596	Sequence 20596, A	233	6	2.2	554	5	PCT-US94-06430-7	Sequence 7, Appl
161	6	2.2	340	4	US-09-252-991A-20599	Sequence 20599, A	234	6	2.2	554	5	PCT-US94-06430-7	Sequence 7, Appl
162	6	2.2	341	3	US-08-928-100-2	Sequence 28871, A	235	6	2.2	564	4	US-09-134-001C-2897	Sequence 2897, Ap
163	6	2.2	341	4	US-08-492-581-2	Sequence 2, Appl	236	6	2.2	578	4	US-09-252-991A-18656	Sequence 18656, A
164	6	2.2	341	4	US-09-425-686-2	Sequence 2, Appl	237	6	2.2	601	4	US-09-252-991A-22594	Sequence 22594, A
165	6	2.2	351	4	US-09-252-991A-31662	Sequence 31662, A	238	6	2.2	602	4	US-09-252-991A-28298	Sequence 28298, A
166	6	2.2	351	4	US-09-252-991A-31662	Sequence 31662, A	239	6	2.2	608	2	US-08-736-770-1	Sequence 1, Appl
167	6	2.2	356	4	US-09-252-991A-32092	Sequence 32092, A	240	6	2.2	615	4	US-09-707-802-9	Sequence 9, Appl
168	6	2.2	358	4	US-09-252-991A-27389	Sequence 27389, A	241	6	2.2	615	4	US-09-707-802-9	Sequence 9, Appl
169	6	2.2	363	4	US-09-328-352-5693	Sequence 5693, Ap	242	6	2.2	615	4	US-09-991-326-9	Sequence 9, Appl
170	6	2.2	375	4	US-09-328-352-5693	Sequence 5693, Ap	243	6	2.2	615	4	US-09-991-326-9	Sequence 9, Appl
171	6	2.2	376	3	US-09-025-691-3	Sequence 7883, Ap	244	6	2.2	632	1	US-08-987-119-2	Sequence 2, Appl
172	6	2.2	376	4	US-09-056-556-502	Sequence 202, Ap	245	6	2.2	653	4	US-09-252-991A-27553	Sequence 27553, A
173	6	2.2	376	4	US-09-072-596-197	Sequence 197, Ap	246	6	2.2	654	4	US-09-252-991A-25801	Sequence 25801, A

247	6	2.2	659	4	US-09-562-737-20	Sequence 20, Appl	320	5	1.8	5	1	US-08-232-815-5	Sequence 5, Appl1
248	6	2.2	661	4	US-09-564-805-240	Sequence 240, App	321	5	1.8	5	1	US-08-350-906-5	Sequence 5, Appl1
249	6	2.2	661	4	US-09-252-991A-18960	Sequence 18960, A	322	5	1.8	5	1	US-07-789-184-126	Sequence 126, App
250	6	2.2	680	4	US-09-107-532A-4798	Sequence 4798, Ap	323	5	1.8	5	1	US-08-704-170-114	Sequence 114, App
251	6	2.2	689	4	US-09-499-964-1	Sequence 1, Appl1	324	5	1.8	5	1	US-08-475-263-126	Sequence 126, App
252	6	2.2	690	4	US-08-935-433-2	Sequence 2, Appl1	325	5	1.8	5	1	US-08-485-886-126	Sequence 126, App
253	6	2.2	690	4	US-09-553-132-2	Sequence 2, Appl1	326	5	1.8	5	2	US-08-477-362-126	Sequence 126, App
254	6	2.2	690	4	US-09-252-991A-32350	Sequence 32350, A	327	5	1.8	5	2	US-08-477-134-126	Sequence 126, App
255	6	2.2	701	4	US-09-252-991A-30741	Sequence 30741, A	328	5	1.8	5	3	US-08-473-489A-126	Sequence 126, App
256	6	2.2	719	4	US-09-252-991A-30052	Sequence 30052, A	329	5	1.8	5	3	US-08-485-695-126	Sequence 126, App
257	6	2.2	721	4	US-09-252-991A-17180	Sequence 17180, A	330	5	1.8	5	3	US-08-018-760-126	Sequence 126, App
258	6	2.2	739	4	US-09-252-991A-32778	Sequence 32778, A	331	5	1.8	5	3	US-09-246-500B-16	Sequence 16, Appl
259	6	2.2	740	1	US-08-309-512-10	Sequence 10, Appl	332	5	1.8	5	5	PCT-US94-02631-114	Sequence 114, App
260	6	2.2	740	5	PCT-US92-08756A-10	Sequence 10, Appl	333	5	1.8	5	5	PCT-US95-04536-5	Sequence 5, Appl1
261	6	2.2	750	3	US-08-202-841A-2	Sequence 2, Appl1	334	5	1.8	6	1	US-08-484-493-9	Sequence 9, Appl1
262	6	2.2	751	4	US-09-402-929-2	Sequence 2, Appl1	335	5	1.8	6	1	US-08-484-494-9	Sequence 9, Appl1
263	6	2.2	752	4	US-09-402-929-6	Sequence 6, Appl1	336	5	1.8	6	2	US-08-345-212-9	Sequence 9, Appl1
264	6	2.2	755	4	US-09-704-611-1	Sequence 1, Appl1	337	5	1.8	6	2	US-08-482-528-160	Sequence 160, App
265	6	2.2	766	5	PCT-US94-00198-6	Sequence 6, Appl1	338	5	1.8	6	3	US-08-482-528-160	Sequence 160, App
266	6	2.2	777	3	US-08-764-870-13	Sequence 13, Appl	339	5	1.8	6	3	US-09-249-003-9	Sequence 9, Appl1
267	6	2.2	777	3	US-08-980-115-13	Sequence 13, Appl	340	5	1.8	6	4	US-09-187-855-2632	Sequence 2632, App
268	6	2.2	784	4	US-09-982-308B-23	Sequence 23, Appl	341	5	1.8	6	4	US-09-187-855-2632	Sequence 2632, App
269	6	2.2	785	4	US-09-134-001C-3866	Sequence 3866, Ap	342	5	1.8	6	4	US-09-187-855-2632	Sequence 2632, App
270	6	2.2	804	4	US-09-134-001C-5218	Sequence 5218, Ap	343	5	1.8	6	4	US-09-685-844-9	Sequence 9, Appl1
271	6	2.2	810	4	US-09-540-824-25	Sequence 25, Appl	344	5	1.8	6	4	US-09-839-542B-2632	Sequence 2632, App
272	6	2.2	830	4	US-09-562-737-34	Sequence 34, Appl	345	5	1.8	6	4	US-09-839-542B-2632	Sequence 2632, App
273	6	2.2	862	3	US-09-315-861-2	Sequence 2, Appl1	346	5	1.8	6	4	US-09-839-542B-2692	Sequence 2692, App
274	6	2.2	862	4	US-09-398-395A-44	Sequence 44, Appl	347	5	1.8	7	1	US-08-232-815-6	Sequence 6, Appl1
275	6	2.2	862	4	US-09-887-586A-44	Sequence 44, Appl	348	5	1.8	7	1	US-08-350-906-6	Sequence 6, Appl1
276	6	2.2	862	4	US-09-895-752-44	Sequence 44, Appl	349	5	1.8	7	1	US-08-308-086-1	Sequence 1, Appl1
277	6	2.2	862	4	US-09-328-352-5527	Sequence 5527, Ap	350	5	1.8	7	4	US-09-187-855-2635	Sequence 2635, App
278	6	2.2	862	4	US-09-903-012B-44	Sequence 44, Ap	351	5	1.8	7	4	US-09-187-855-2635	Sequence 2635, App
279	6	2.2	886	4	US-09-328-352-6985	Sequence 6985, Ap	352	5	1.8	7	4	US-09-187-855-2635	Sequence 2635, App
280	6	2.2	905	4	US-09-134-001C-3782	Sequence 3782, Ap	353	5	1.8	7	4	US-09-839-542B-2635	Sequence 2635, App
281	6	2.2	909	3	US-08-936-135-8	Sequence 8, Appl1	354	5	1.8	7	4	US-09-839-542B-2635	Sequence 2635, App
282	6	2.2	909	3	US-08-936-135-10	Sequence 10, Appl	355	5	1.8	7	4	US-09-839-542B-2693	Sequence 2693, App
283	6	2.2	909	3	US-08-936-135-18	Sequence 18, Appl	356	5	1.8	7	5	PCT-US95-04536-6	Sequence 6, Appl1
284	6	2.2	914	3	US-08-936-135-12	Sequence 12, Appl	357	5	1.8	7	6	5342922-23	Patent No. 5342922
285	6	2.2	925	4	US-09-116-473-2	Sequence 2, Appl1	358	5	1.8	8	2	US-08-616-669A-7	Sequence 7, Appl1
286	6	2.2	926	3	US-08-936-135-14	Sequence 14, Appl	359	5	1.8	8	3	US-09-326-440-7	Sequence 7, Appl1
287	6	2.2	926	3	US-08-936-135-20	Sequence 20, Appl	360	5	1.8	8	4	US-09-187-855-2638	Sequence 2638, App
288	6	2.2	931	3	US-08-936-135-16	Sequence 16, Appl	361	5	1.8	8	4	US-09-187-855-2638	Sequence 2638, App
289	6	2.2	932	4	US-09-328-352-7453	Sequence 7453, Ap	362	5	1.8	8	4	US-09-187-855-2638	Sequence 2638, App
290	6	2.2	933	3	US-08-764-870-14	Sequence 14, Appl	363	5	1.8	8	4	US-09-839-542B-2638	Sequence 2638, App
291	6	2.2	933	4	US-08-980-115-14	Sequence 14, Appl	364	5	1.8	8	4	US-09-839-542B-2664	Sequence 2664, App
292	6	2.2	940	4	US-09-198-452A-111	Sequence 111, App	365	5	1.8	8	4	US-09-839-542B-2694	Sequence 2694, App
293	6	2.2	958	1	US-08-426-236-4	Sequence 4, Appl1	366	5	1.8	9	2	US-08-704-655-1	Sequence 1, Appl1
294	6	2.2	990	4	US-09-627-376-7	Sequence 7, Appl1	367	5	1.8	9	2	US-08-704-655-16	Sequence 16, Appl1
295	6	2.2	1003	4	US-09-521-511C-11	Sequence 11, Appl	368	5	1.8	9	2	US-08-704-655-17	Sequence 17, Appl
296	6	2.2	1006	4	US-09-023-905A-12	Sequence 12, Appl	369	5	1.8	9	2	US-08-704-655-19	Sequence 19, Appl
297	6	2.2	1013	4	US-09-062-126-3	Sequence 3, Appl1	370	5	1.8	9	2	US-08-704-655-21	Sequence 21, Appl
298	6	2.2	1022	3	US-08-772-270A-2	Sequence 2, Appl1	371	5	1.8	9	2	US-08-704-655-21	Sequence 21, Appl
299	6	2.2	1039	6	US-08-665-574C-16	Sequence 16, Appl	372	5	1.8	9	4	US-08-630-915A-59	Sequence 59, Appl
300	6	2.2	1112	4	US-09-252-991A-27256	Sequence 27256, A	373	5	1.8	9	4	US-09-187-855-2641	Sequence 2641, App
301	6	2.2	1113	3	US-09-629-616-3	Sequence 3, Appl1	374	5	1.8	9	4	US-09-839-542B-2641	Sequence 2641, App
302	6	2.2	1138	4	US-09-572-191-2	Sequence 2, Appl1	375	5	1.8	10	1	US-08-089-994A-18	Sequence 18, Appl
303	6	2.2	1388	4	US-09-723-262-2	Sequence 2, Appl1	376	5	1.8	10	1	US-08-444-618-9	Sequence 9, Appl1
304	6	2.2	1388	4	US-09-723-219-2	Sequence 2, Appl1	377	5	1.8	10	1	US-08-199-776-13	Sequence 13, Appl
305	6	2.2	1501	2	US-08-447-464-3	Sequence 3, Appl1	378	5	1.8	10	1	US-08-199-776-14	Sequence 14, Appl
306	6	2.2	1501	2	US-08-716-679-3	Sequence 3, Appl1	379	5	1.8	10	1	US-08-199-776-15	Sequence 15, Appl
307	6	2.2	1911	1	US-08-348-006B-5	Sequence 5, Appl1	380	5	1.8	10	1	US-08-199-776-16	Sequence 16, Appl
308	6	2.2	1911	2	US-08-800-825A-5	Sequence 5, Appl1	381	5	1.8	10	1	US-08-426-311-3	Sequence 3, Appl1
309	6	2.2	1911	3	US-09-158-657-5	Sequence 5, Appl1	382	5	1.8	10	2	US-08-623-833B-21	Sequence 21, Appl
310	6	2.2	1911	5	PCT-US94-10166-5	Sequence 5, Appl1	383	5	1.8	10	3	US-08-663-731-13	Sequence 13, Appl
311	6	2.2	2071	3	US-09-415-522-6	Sequence 6, Appl1	384	5	1.8	10	3	US-08-663-731-14	Sequence 14, Appl
312	6	2.2	2233	2	US-08-569-853-1	Sequence 1, Appl1	385	5	1.8	10	3	US-08-663-731-15	Sequence 15, Appl
313	6	2.2	2233	2	US-08-569-853-2	Sequence 2, Appl1	386	5	1.8	10	3	US-08-663-731-16	Sequence 16, Appl
314	6	2.2	2233	3	US-08-987-439-1	Sequence 1, Appl1	387	5	1.8	10	3	US-08-879-338-13	Sequence 13, Appl
315	6	2.2	2504	4	US-09-328-352-5821	Sequence 5821, Ap	388	5	1.8	10	3	US-08-879-338-14	Sequence 14, Appl
316	6	2.2	2713	5	PCT-US96-01735-1	Sequence 1, Appl1	389	5	1.8	10	3	US-08-879-338-15	Sequence 15, Appl
317	6	2.2	2749	5	US-09-385-222A-4	Sequence 4, Appl1	390	5	1.8	10	3	US-08-879-338-16	Sequence 16, Appl
318	6	2.2	3060	2	US-08-487-826B-14	Sequence 14, Appl	391	5	1.8	10	4	US-09-643-597-229	Sequence 229, App
319	6	2.2	3562	4	US-09-679-279-14	Sequence 14, Appl	392	5	1.8	10	4	US-09-293-238B-13	Sequence 13, Appl

393	5	1.8	10	4	US-09-293-238B-14	Sequence 14, Appl	466	5	1.8	16	1	US-08-037-574A-95	Sequence 95, Appl
394	5	1.8	10	4	US-09-293-238B-15	Sequence 15, Appl	467	5	1.8	16	1	US-08-037-574A-96	Sequence 96, Appl
395	5	1.8	10	4	US-09-293-238B-16	Sequence 16, Appl	468	5	1.8	16	1	US-08-218-329-28	Sequence 28, Appl
396	5	1.8	10	4	US-09-480-884A-229	Sequence 229, App	469	5	1.8	16	1	US-08-218-329-29	Sequence 29, Appl
397	5	1.8	10	4	US-09-542-615A-229	Sequence 229, App	470	5	1.8	16	1	US-08-218-329-30	Sequence 30, Appl
398	5	1.8	10	4	US-09-606-421B-229	Sequence 229, App	471	5	1.8	16	1	US-08-218-329-31	Sequence 31, Appl
399	5	1.8	10	4	PCT-US94-07605-18	Sequence 18, Appl	472	5	1.8	16	1	US-08-218-329-32	Sequence 32, Appl
400	5	1.8	10	5	PCT-US95-02044-13	Sequence 13, Appl	473	5	1.8	16	1	US-08-218-329-33	Sequence 33, Appl
401	5	1.8	10	5	PCT-US95-02044-14	Sequence 14, Appl	474	5	1.8	16	1	US-08-218-329-34	Sequence 34, Appl
402	5	1.8	10	5	PCT-US95-02044-15	Sequence 15, Appl	475	5	1.8	16	1	US-08-218-329-35	Sequence 35, Appl
403	5	1.8	10	5	PCT-US95-02044-16	Sequence 16, Appl	476	5	1.8	16	1	US-08-218-329-36	Sequence 36, Appl
404	5	1.8	10	6	5223391-2	Patent No. 5223391	477	5	1.8	16	1	US-08-218-329-37	Sequence 37, Appl
405	5	1.8	11	1	US-08-089-994A-17	Sequence 17, Appl	478	5	1.8	16	1	US-08-218-329-93	Sequence 93, Appl
406	5	1.8	11	1	US-08-443-568B-19	Sequence 19, Appl	479	5	1.8	16	1	US-08-218-329-94	Sequence 94, Appl
407	5	1.8	11	2	US-08-823-932-3	Sequence 3, Appl	480	5	1.8	16	1	US-08-218-329-95	Sequence 95, Appl
408	5	1.8	11	3	US-09-181-816-3	Sequence 3, Appl	481	5	1.8	16	1	PCT-US91-09422-22	Sequence 22, Appl
409	5	1.8	11	3	US-09-181-816-6	Sequence 6, Appl	482	5	1.8	17	1	US-08-037-574A-38	Sequence 38, Appl
410	5	1.8	11	5	PCT-US94-06997-19	Sequence 19, Appl	483	5	1.8	17	1	US-08-037-574A-39	Sequence 39, Appl
411	5	1.8	11	5	PCT-US94-07605-17	Sequence 17, Appl	484	5	1.8	17	1	US-08-037-574A-40	Sequence 40, Appl
412	5	1.8	11	5	PCT-US95-12987-9	Sequence 9, Appl	485	5	1.8	17	1	US-08-037-574A-41	Sequence 41, Appl
413	5	1.8	12	1	US-08-199-776-22	Sequence 22, Appl	486	5	1.8	17	1	US-08-037-574A-42	Sequence 42, Appl
414	5	1.8	12	3	US-08-663-731-22	Sequence 22, Appl	487	5	1.8	17	1	US-08-037-574A-43	Sequence 43, Appl
415	5	1.8	12	3	US-08-879-338-22	Sequence 22, Appl	488	5	1.8	17	1	US-08-037-574A-44	Sequence 44, Appl
416	5	1.8	12	3	US-08-747-599A-21	Sequence 21, Appl	489	5	1.8	17	1	US-08-037-574A-45	Sequence 45, Appl
417	5	1.8	12	3	US-08-891-640-6	Sequence 6, Appl	490	5	1.8	17	1	US-08-037-574A-46	Sequence 46, Appl
418	5	1.8	12	4	US-09-293-238B-22	Sequence 22, Appl	491	5	1.8	17	1	US-08-037-574A-47	Sequence 47, Appl
419	5	1.8	12	5	PCT-US93-05701-3	Sequence 3, Appl	492	5	1.8	17	1	US-08-218-329-43	Sequence 43, Appl
420	5	1.8	12	5	PCT-US95-02044-22	Sequence 22, Appl	493	5	1.8	17	1	US-08-218-329-38	Sequence 38, Appl
421	5	1.8	13	2	US-08-723-415B-16	Sequence 16, Appl	494	5	1.8	17	1	US-08-218-329-39	Sequence 39, Appl
422	5	1.8	13	3	US-09-189-627A-16	Sequence 16, Appl	495	5	1.8	17	1	US-08-218-329-40	Sequence 40, Appl
423	5	1.8	13	4	US-09-710-861-16	Sequence 16, Appl	496	5	1.8	17	1	US-08-218-329-41	Sequence 41, Appl
424	5	1.8	13	6	5168051-13	Patent No. 5168051	497	5	1.8	17	1	US-08-218-329-42	Sequence 42, Appl
425	5	1.8	14	1	US-08-432-617-14	Sequence 14, Appl	498	5	1.8	17	1	US-08-218-329-43	Sequence 43, Appl
426	5	1.8	14	1	US-08-199-776-18	Sequence 18, Appl	499	5	1.8	17	1	US-08-218-329-44	Sequence 44, Appl
427	5	1.8	14	3	US-08-663-731-18	Sequence 18, Appl	500	5	1.8	17	1	US-08-218-329-45	Sequence 45, Appl
428	5	1.8	14	3	US-08-879-338-18	Sequence 18, Appl	501	5	1.8	17	1	US-08-218-329-46	Sequence 46, Appl
429	5	1.8	14	3	US-09-195-868-27	Sequence 27, Appl	502	5	1.8	17	1	US-08-218-329-47	Sequence 47, Appl
430	5	1.8	14	4	US-09-293-238B-18	Sequence 18, Appl	503	5	1.8	17	1	US-08-218-329-60	Sequence 60, Appl
431	5	1.8	14	5	PCT-US93-06751-106	Sequence 106, App	504	5	1.8	17	1	US-08-379-613-1	Sequence 1, Appl
432	5	1.8	14	5	PCT-US95-02044-18	Sequence 18, Appl	505	5	1.8	17	1	US-08-379-613-2	Sequence 2, Appl
433	5	1.8	15	1	US-08-199-776-20	Sequence 20, Appl	506	5	1.8	17	1	US-08-774-325-2	Sequence 2, Appl
434	5	1.8	15	1	US-08-488-212A-31	Sequence 31, Appl	507	5	1.8	17	2	US-08-774-325-3	Sequence 3, Appl
435	5	1.8	15	2	US-08-320-306-31	Sequence 31, Appl	508	5	1.8	17	2	US-09-115-209-63	Sequence 63, Appl
436	5	1.8	15	2	US-08-488-209B-31	Sequence 31, Appl	509	5	1.8	17	2	US-09-115-209-64	Sequence 64, Appl
437	5	1.8	15	2	US-08-408-011-31	Sequence 31, Appl	510	5	1.8	17	2	US-09-115-209-65	Sequence 65, Appl
438	5	1.8	15	2	US-08-484-905-47	Sequence 47, Appl	511	5	1.8	17	2	US-09-115-209-66	Sequence 66, Appl
439	5	1.8	15	2	US-08-466-860-14	Sequence 14, Appl	512	5	1.8	17	3	US-08-488-551B-801	Sequence 801, App
440	5	1.8	15	3	US-08-481-985B-47	Sequence 47, Appl	513	5	1.8	17	3	US-08-860-300-10	Sequence 10, Appl
441	5	1.8	15	3	US-08-663-731-20	Sequence 20, Appl	514	5	1.8	17	3	US-09-269-062-1	Sequence 1, Appl
442	5	1.8	15	3	US-08-879-338-20	Sequence 20, Appl	515	5	1.8	17	4	US-09-887-439-1	Sequence 1, Appl
443	5	1.8	15	3	US-08-472-040A-14	Sequence 14, Appl	516	5	1.8	17	4	US-09-885-932-1	Sequence 1, Appl
444	5	1.8	15	3	US-08-370-476-47	Sequence 47, Appl	517	5	1.8	17	4	US-09-416-509C-4	Sequence 4, Appl
445	5	1.8	15	3	US-08-276-776-14	Sequence 14, Appl	518	5	1.8	17	5	PCT-US91-09422-21	Sequence 21, Appl
446	5	1.8	15	3	US-08-471-209-14	Sequence 14, Appl	519	5	1.8	17	5	PCT-US95-16596-10	Sequence 10, Appl
447	5	1.8	15	3	US-08-743-168B-25	Sequence 25, Appl	520	5	1.8	17	5	US-08-218-025A-116	Sequence 116, App
448	5	1.8	15	4	US-08-992-877-27	Sequence 27, Appl	521	5	1.8	18	1	US-08-295-085-10	Sequence 10, Appl
449	5	1.8	15	4	US-09-053-611-36	Sequence 36, Appl	522	5	1.8	18	1	US-07-800-364B-2	Sequence 2, Appl
450	5	1.8	15	4	US-09-293-238B-20	Sequence 20, Appl	523	5	1.8	18	3	US-08-854-764-7	Sequence 7, Appl
451	5	1.8	15	5	PCT-US95-02044-20	Sequence 20, Appl	524	5	1.8	18	3	US-08-990-888-38	Sequence 38, Appl
452	5	1.8	15	5	PCT-US96-10435-25	Sequence 25, Appl	525	5	1.8	18	4	PCT-US91-03388-2	Sequence 2, Appl
453	5	1.8	16	1	US-08-218-025A-111	Sequence 111, App	526	5	1.8	18	5	PCT-US95-09377-7	Sequence 7, Appl
454	5	1.8	16	1	US-08-037-574A-33	Sequence 33, Appl	527	5	1.8	18	5	PCT-US95-10741-10	Sequence 10, Appl
455	5	1.8	16	1	US-08-037-574A-29	Sequence 29, Appl	528	5	1.8	19	1	US-08-335-198-10	Sequence 10, Appl
456	5	1.8	16	1	US-08-037-574A-30	Sequence 30, Appl	529	5	1.8	19	2	US-08-704-655-28	Sequence 28, Appl
457	5	1.8	16	1	US-08-037-574A-31	Sequence 31, Appl	530	5	1.8	19	3	US-08-460-040-3	Sequence 3, Appl
458	5	1.8	16	1	US-08-037-574A-32	Sequence 32, Appl	531	5	1.8	19	3	US-08-781-420-4	Sequence 4, Appl
459	5	1.8	16	1	US-08-037-574A-33	Sequence 33, Appl	532	5	1.8	19	4	US-08-874-102-4	Sequence 4, Appl
460	5	1.8	16	1	US-08-037-574A-34	Sequence 34, Appl	533	5	1.8	19	4	US-08-984-919A-4	Sequence 4, Appl
461	5	1.8	16	1	US-08-037-574A-35	Sequence 35, Appl	534	5	1.8	19	4	US-09-006-595A-4	Sequence 4, Appl
462	5	1.8	16	1	US-08-037-574A-36	Sequence 36, Appl	535	5	1.8	19	4	US-09-149-476-599	Sequence 599, App
463	5	1.8	16	1	US-08-037-574A-37	Sequence 37, Appl	536	5	1.8	19	4	US-09-081-975-13	Sequence 13, Appl
464	5	1.8	16	1	US-08-037-574A-59	Sequence 59, Appl	537	5	1.8	19	4	US-08-218-025A-53	Sequence 53, Appl
465	5	1.8	16	1	US-08-037-574A-94	Sequence 94, Appl	538	5	1.8	20	1		

539	5	1.8	20	1	US-08-218-025A-54	Sequence 54, App1	612	5	1.8	26	3	US-08-965-762-35	Sequence 35, App1
540	5	1.8	20	1	US-08-199-776-21	Sequence 21, App1	613	5	1.8	26	4	US-09-227-357-346	Sequence 346, App
541	5	1.8	20	2	US-08-482-142-9	Sequence 9, App1	614	5	1.8	26	4	US-09-315-926A-77	Sequence 77, App1
542	5	1.8	20	2	US-07-952-853-20	Sequence 20, App1	615	5	1.8	27	6	5466783-26	Patent No. 5466783
543	5	1.8	20	2	US-08-478-572-9	Sequence 9, App1	616	5	1.8	27	6	5498694-8	Patent No. 5498694
544	5	1.8	20	2	US-08-914-848-20	Sequence 20, App1	617	5	1.8	28	2	US-08-482-142-27	Sequence 27, App1
545	5	1.8	20	3	US-08-663-731-21	Sequence 21, App1	618	5	1.8	28	2	US-08-478-572-27	Sequence 27, App1
546	5	1.8	20	3	US-08-879-338-21	Sequence 21, App1	619	5	1.8	28	3	US-09-040-485-3	Sequence 3, App1
547	5	1.8	20	3	US-08-602-999A-114	Sequence 114, App	620	5	1.8	28	3	US-08-484-296-27	Sequence 27, App1
548	5	1.8	20	3	US-08-484-296-9	Sequence 9, App1	621	5	1.8	28	4	US-09-227-357-477	Sequence 477, App
549	5	1.8	20	4	US-09-500-124-114	Sequence 114, App	622	5	1.8	28	4	US-09-205-258-638	Sequence 638, App
550	5	1.8	20	4	US-09-293-238B-21	Sequence 21, App1	623	5	1.8	28	6	5168045-3	Patent No. 5168045
551	5	1.8	20	4	US-09-736-457-1855	Sequence 1855, Ap	624	5	1.8	29	4	US-09-205-258-1187	Sequence 1187, Ap
552	5	1.8	20	4	US-09-736-457-1856	Sequence 1856, Ap	625	5	1.8	30	1	US-08-145-708A-19	Sequence 19, App1
553	5	1.8	20	4	US-09-319-570A-5	Sequence 5, App1	626	5	1.8	30	1	US-08-087-772A-9	Sequence 9, App1
554	5	1.8	20	5	PCT-US95-02044-21	Sequence 21, App1	627	5	1.8	30	2	US-08-331-454-19	Sequence 19, App1
555	5	1.8	20	5	PCT-US95-04481-1	Sequence 1, App1	628	5	1.8	30	3	US-08-818-112-128	Sequence 128, App
556	5	1.8	21	1	US-07-768-269-47	Sequence 47, App1	629	5	1.8	30	4	US-08-818-111-123	Sequence 123, App
557	5	1.8	21	1	US-08-199-776-5	Sequence 5, App1	630	5	1.8	30	4	US-09-056-556-128	Sequence 128, App
558	5	1.8	21	1	US-08-306-871-44	Sequence 44, App1	631	5	1.8	30	4	US-09-072-596-123	Sequence 123, App
559	5	1.8	21	1	US-08-569-959-44	Sequence 44, App1	632	5	1.8	30	6	5177189-1	Patent No. 5177189
560	5	1.8	21	3	US-09-082-737-7	Sequence 7, App1	633	5	1.8	30	6	5182364-3	Patent No. 5182364
561	5	1.8	21	3	US-08-075-520A-37	Sequence 37, App1	634	5	1.8	31	2	US-08-845-926-3	Sequence 3, App1
562	5	1.8	21	3	US-08-663-731-5	Sequence 5, App1	635	5	1.8	31	4	US-09-351-296-3	Sequence 3, App1
563	5	1.8	21	3	US-08-879-338-5	Sequence 5, App1	636	5	1.8	32	1	US-07-916-901-9	Sequence 9, App1
564	5	1.8	21	3	US-08-879-338-26	Sequence 26, App1	637	5	1.8	32	1	US-08-214-770-2	Sequence 2, App1
565	5	1.8	21	3	US-08-652-877-61	Sequence 61, App1	638	5	1.8	32	1	US-08-440-103-35	Sequence 35, App1
566	5	1.8	21	3	US-08-476-515A-61	Sequence 61, App1	639	5	1.8	32	1	US-08-440-542-35	Sequence 35, App1
567	5	1.8	21	4	US-09-293-238B-5	Sequence 5, App1	640	5	1.8	32	1	US-08-231-368-35	Sequence 35, App1
568	5	1.8	21	4	US-09-293-238B-26	Sequence 26, App1	641	5	1.8	32	1	US-08-440-210-35	Sequence 35, App1
569	5	1.8	21	5	PCT-US95-02044-5	Sequence 5, App1	642	5	1.8	32	3	US-08-966-269-5	Sequence 5, App1
570	5	1.8	22	1	US-07-933-469A-7	Sequence 7, App1	643	5	1.8	32	4	US-09-046-604-35	Sequence 35, App1
571	5	1.8	22	1	US-08-250-310-7	Sequence 7, App1	644	5	1.8	32	4	US-09-436-183A-5	Sequence 5, App1
572	5	1.8	22	1	US-08-469-615-11	Sequence 11, App1	645	5	1.8	32	5	PCT-US95-02885-2	Sequence 2, App1
573	5	1.8	22	1	US-08-439-404-7	Sequence 7, App1	646	5	1.8	33	1	US-08-520-599-2	Sequence 2, App1
574	5	1.8	22	1	US-08-466-763-11	Sequence 11, App1	647	5	1.8	33	4	US-09-402-664A-15	Sequence 15, App1
575	5	1.8	22	2	US-08-411-142A-11	Sequence 11, App1	648	5	1.8	33	4	US-09-205-258-865	Sequence 865, App
576	5	1.8	22	3	US-08-839-327-7	Sequence 7, App1	649	5	1.8	34	1	US-07-915-247A-5	Sequence 5, App1
577	5	1.8	22	3	US-09-340-781B-7	Sequence 7, App1	650	5	1.8	34	1	US-08-443-863-5	Sequence 5, App1
578	5	1.8	22	3	US-08-557-006C-34	Sequence 34, App1	651	5	1.8	34	1	US-08-448-070-5	Sequence 5, App1
579	5	1.8	22	4	US-09-354-231B-48	Sequence 48, App1	652	5	1.8	34	1	US-08-449-500-5	Sequence 5, App1
580	5	1.8	22	4	US-09-128-602B-48	Sequence 48, App1	653	5	1.8	34	1	US-08-449-500-65	Sequence 65, App1
581	5	1.8	22	6	5427926-1	Patent No. 5427926	654	5	1.8	34	1	US-08-449-317A-5	Sequence 5, App1
582	5	1.8	23	1	US-07-778-233B-51	Sequence 51, App1	655	5	1.8	34	1	US-08-449-317A-65	Sequence 65, App1
583	5	1.8	23	1	US-07-963-321-51	Sequence 51, App1	656	5	1.8	34	2	US-08-477-022-5	Sequence 5, App1
584	5	1.8	23	1	US-08-290-641-51	Sequence 51, App1	657	5	1.8	34	2	US-08-477-022-65	Sequence 65, App1
585	5	1.8	23	1	US-08-340-428B-41	Sequence 41, App1	658	5	1.8	34	2	US-08-449-447-65	Sequence 65, App1
586	5	1.8	23	1	US-08-548-540-51	Sequence 51, App1	659	5	1.8	34	2	US-08-449-447-65	Sequence 65, App1
587	5	1.8	23	1	US-08-189-331-180	Sequence 180, App	660	5	1.8	34	2	US-08-445-520B-6	Sequence 6, App1
588	5	1.8	23	2	US-08-471-068-180	Sequence 180, App	661	5	1.8	34	2	US-08-184-328-5	Sequence 5, App1
589	5	1.8	23	3	US-08-905-223-382	Sequence 382, App	662	5	1.8	34	2	US-08-184-328-65	Sequence 65, App1
590	5	1.8	23	4	US-09-645-470-21	Sequence 21, App1	663	5	1.8	34	2	US-08-521-097-5	Sequence 5, App1
591	5	1.8	23	4	US-09-645-470-27	Sequence 27, App1	664	5	1.8	34	2	US-08-521-097-65	Sequence 65, App1
592	5	1.8	23	5	PCT-US93-07306-41	Sequence 41, App1	665	5	1.8	34	3	US-09-239-864A-6	Sequence 6, App1
593	5	1.8	23	5	PCT-US96-09809-51	Sequence 51, App1	666	5	1.8	35	1	US-08-487-890A-69	Sequence 69, App1
594	5	1.8	24	2	US-08-482-142-28	Sequence 28, App1	667	5	1.8	35	1	US-08-142-551B-111	Sequence 111, App
595	5	1.8	24	2	US-08-146-028-72	Sequence 72, App1	668	5	1.8	35	2	US-08-478-435-69	Sequence 69, App1
596	5	1.8	24	2	US-08-478-572-28	Sequence 72, App1	669	5	1.8	35	2	US-08-337-483-69	Sequence 69, App1
597	5	1.8	24	2	US-08-723-425A-72	Sequence 72, App1	670	5	1.8	35	2	US-08-478-373-69	Sequence 69, App1
598	5	1.8	24	3	US-09-112-206-78	Sequence 72, App1	671	5	1.8	35	3	US-08-474-671-69	Sequence 69, App1
599	5	1.8	24	3	US-08-484-296-28	Sequence 28, App1	672	5	1.8	35	3	US-08-483-577A-69	Sequence 69, App1
600	5	1.8	24	5	PCT-US95-04481-27	Sequence 27, App1	673	5	1.8	35	3	US-08-897-438-69	Sequence 69, App1
601	5	1.8	24	6	5204096-10	Patent No. 5204096	674	5	1.8	35	4	US-08-637-654-69	Sequence 69, App1
602	5	1.8	25	2	US-08-482-142-120	Sequence 120, App	675	5	1.8	35	4	US-08-649-518-69	Sequence 69, App1
603	5	1.8	25	2	US-08-493-235-20	Sequence 20, App1	676	5	1.8	36	2	US-08-146-028-71	Sequence 71, App1
604	5	1.8	25	2	US-08-478-572-120	Sequence 120, App	677	5	1.8	36	2	US-08-146-028-127	Sequence 127, App
605	5	1.8	25	3	US-08-484-296-120	Sequence 120, App	678	5	1.8	36	3	US-08-554-385-23	Sequence 23, App1
606	5	1.8	25	4	US-09-911-827-35	Sequence 35, App1	679	5	1.8	36	3	US-08-723-425A-71	Sequence 71, App1
607	5	1.8	25	4	US-09-911-882-35	Sequence 35, App1	680	5	1.8	36	3	US-08-723-425A-127	Sequence 127, App
608	5	1.8	25	4	US-09-911-888-35	Sequence 35, App1	681	5	1.8	36	3	US-09-112-206-71	Sequence 71, App1
609	5	1.8	26	2	US-08-146-028-73	Sequence 73, App1	682	5	1.8	36	3	US-09-112-206-127	Sequence 127, App
610	5	1.8	26	3	US-08-723-425A-73	Sequence 73, App1	683	5	1.8	39	3	US-08-973-462-23	Sequence 23, App1
611	5	1.8	26	3	US-09-112-206-73	Sequence 73, App1	684	5	1.8	40	3	US-08-756-849-124	Sequence 124, App

685	5	1.8	40	4	US-08-990-888-7	Sequence 7, Appl1	758	5	1.8	60	4	US-09-107-532A-3656	Sequence 3656, Ap
686	5	1.8	41	1	US-08-865-773-8	Sequence 8, Appl1	759	5	1.8	61	4	US-09-107-532A-4373	Sequence 4373, Ap
687	5	1.8	41	1	US-08-865-773-9	Sequence 9, Appl1	760	5	1.8	62	2	US-08-800-264A-12	Sequence 12, Appl
688	5	1.8	41	3	US-08-981-189B-4	Sequence 4, Appl1	761	5	1.8	62	2	US-09-018-628-12	Sequence 12, Appl
689	5	1.8	41	3	US-08-981-189B-5	Sequence 5, Appl1	762	5	1.8	62	3	US-09-273-378-12	Sequence 12, Appl
690	5	1.8	41	3	US-09-260-846-25	Sequence 25, Appl	763	5	1.8	62	3	US-09-409-605-12	Sequence 12, Appl
691	5	1.8	41	4	US-09-400-716-6	Sequence 6, Appl1	764	5	1.8	62	4	US-09-134-001C-3788	Sequence 3788, Ap
692	5	1.8	41	4	US-09-400-716-7	Sequence 7, Appl1	765	5	1.8	62	4	US-09-107-532A-5441	Sequence 5441, Ap
693	5	1.8	41	4	US-09-424-127-8	Sequence 8, Appl1	766	5	1.8	62	4	US-09-107-532A-6676	Sequence 6676, Ap
694	5	1.8	41	4	US-09-424-127-9	Sequence 9, Appl1	767	5	1.8	63	4	US-09-072-596-284	Sequence 284, App
695	5	1.8	41	6	5177060-6	Patent No. 5177060	768	5	1.8	64	4	US-09-695-458-31	Sequence 31, Appl
696	5	1.8	41	6	5177060-7	Patent No. 5177060	769	5	1.8	64	4	US-09-252-991A-23479	Sequence 23479, A
697	5	1.8	42	1	US-08-520-599-1	Sequence 1, Appl1	770	5	1.8	65	1	US-08-358-160-149	Sequence 149, App
698	5	1.8	42	2	US-08-766-858A-31	Sequence 31, Appl	771	5	1.8	65	2	US-08-633-879C-16	Sequence 16, Appl
699	5	1.8	43	2	US-08-499-676A-29	Sequence 29, Appl	772	5	1.8	65	4	US-09-695-458-38	Sequence 38, Appl
700	5	1.8	43	2	US-08-499-676A-30	Sequence 30, Appl	773	5	1.8	65	4	US-09-252-991A-23587	Sequence 23587, A
701	5	1.8	43	6	US-08-499-676A-31	Sequence 31, Appl	774	5	1.8	65	4	US-09-107-532A-7115	Sequence 7115, Ap
702	5	1.8	43	6	5185440-6	Patent No. 5185440	775	5	1.8	66	3	US-09-135-020-110	Sequence 110, App
703	5	1.8	44	3	US-08-961-083-88	Sequence 88, Appl	776	5	1.8	66	3	US-09-135-010A-110	Sequence 110, App
704	5	1.8	44	4	US-08-990-888-2	Sequence 2, Appl1	777	5	1.8	66	4	US-09-444-871-110	Sequence 110, App
705	5	1.8	44	4	US-09-288-143-112	Sequence 112, App	778	5	1.8	66	4	US-09-597-735-110	Sequence 110, App
706	5	1.8	44	4	US-09-536-784-98	Sequence 88, Appl	779	5	1.8	66	4	US-09-444-295-110	Sequence 110, App
707	5	1.8	45	1	US-08-056-200-102	Sequence 102, App	780	5	1.8	66	4	US-09-597-732-110	Sequence 110, App
708	5	1.8	45	1	US-08-056-200-106	Sequence 106, App	781	5	1.8	66	4	US-09-597-731-110	Sequence 110, App
709	5	1.8	45	1	US-08-062-472B-24	Sequence 24, Appl	782	5	1.8	67	4	US-09-134-001C-5270	Sequence 5270, Ap
710	5	1.8	45	2	US-08-637-759B-236	Sequence 236, App	783	5	1.8	67	4	US-09-328-352-4211	Sequence 4211, Ap
711	5	1.8	45	2	US-08-800-644-102	Sequence 102, App	784	5	1.8	68	1	US-07-756-250-10	Sequence 10, Appl
712	5	1.8	45	2	US-08-800-644-106	Sequence 106, App	785	5	1.8	68	1	US-08-606-789-2	Sequence 2, Appl1
713	5	1.8	45	3	US-08-871-355A-236	Sequence 236, App	786	5	1.8	68	1	US-08-606-789-4	Sequence 4, Appl1
714	5	1.8	45	3	US-09-201-945-236	Sequence 236, App	787	5	1.8	68	2	US-09-111-348-2	Sequence 2, Appl1
715	5	1.8	45	4	US-09-227-357-345	Sequence 345, App	788	5	1.8	68	2	US-09-111-348-4	Sequence 4, Appl1
716	5	1.8	47	1	US-08-222-616-10	Sequence 10, Appl	789	5	1.8	69	5	US-09-107-532A-6222	Sequence 6222, Ap
717	5	1.8	47	1	US-08-014-153D-20	Sequence 20, Appl	790	5	1.8	71	1	PCT-US95-06406A-5	Sequence 5, Appl1
718	5	1.8	47	4	US-08-446-648-10	Sequence 10, Appl	791	5	1.8	71	1	US-08-306-871-28	Sequence 28, Appl
719	5	1.8	47	4	US-08-858-207A-487	Sequence 487, App	792	5	1.8	71	1	US-08-569-959-28	Sequence 28, Appl
720	5	1.8	47	4	US-09-149-476-364	Sequence 364, App	793	5	1.8	71	3	US-08-297-395-8	Sequence 8, Appl1
721	5	1.8	47	5	PCT-US95-04228-10	Sequence 10, Appl	794	5	1.8	71	4	US-08-858-207A-488	Sequence 488, App
722	5	1.8	48	2	US-08-469-537A-62	Sequence 62, Appl	795	5	1.8	72	4	US-09-107-532A-3307	Sequence 3307, Ap
723	5	1.8	48	2	US-08-469-537A-69	Sequence 69, Appl	796	5	1.8	72	4	US-09-107-532A-6789	Sequence 6789, Ap
724	5	1.8	48	2	US-08-683-262B-67	Sequence 67, Appl	797	5	1.8	73	4	US-09-107-532A-4354	Sequence 4354, Ap
725	5	1.8	48	3	US-09-361-707-67	Sequence 67, Appl	798	5	1.8	73	4	US-09-107-532A-4731	Sequence 4731, Ap
726	5	1.8	48	3	US-08-891-640-5	Sequence 5, Appl1	799	5	1.8	74	3	US-09-177-249-56	Sequence 56, Appl
727	5	1.8	48	4	US-09-227-357-241	Sequence 241, App	800	5	1.8	74	4	US-09-328-352-6733	Sequence 6733, Ap
728	5	1.8	48	6	5168045-6	Patent No. 5168045	801	5	1.8	75	4	US-09-254-352B-30	Sequence 30, Appl
729	5	1.8	48	6	5182364-14	Patent No. 5182364	802	5	1.8	75	4	US-09-105-058C-10	Sequence 10, Appl
730	5	1.8	49	1	PCT-US91-03588-8	Sequence 8, Appl1	803	5	1.8	76	3	US-07-741-453A-14	Sequence 14, Appl
731	5	1.8	49	5	US-08-685-871-64	Sequence 64, Appl	804	5	1.8	76	4	US-09-107-532A-3890	Sequence 3890, Ap
732	5	1.8	50	3	US-08-519-103-16	Sequence 16, Appl	805	5	1.8	77	4	US-09-483-846B-1	Sequence 1, Appl1
733	5	1.8	51	1	US-08-519-103-17	Sequence 17, Appl	806	5	1.8	77	4	US-08-311-731A-223	Sequence 223, Appl
734	5	1.8	51	1	US-08-519-103-18	Sequence 18, Appl	807	5	1.8	77	4	US-09-107-532A-5894	Sequence 5894, Ap
735	5	1.8	51	3	US-09-018-635-16	Sequence 16, Appl	808	5	1.8	78	3	US-09-100-804A-29	Sequence 29, Appl
736	5	1.8	51	3	US-09-018-635-17	Sequence 17, Appl	809	5	1.8	78	4	US-08-675-499A-24	Sequence 24, Appl
737	5	1.8	51	3	US-09-018-635-18	Sequence 18, Appl	810	5	1.8	78	4	US-08-675-499A-27	Sequence 27, Appl
738	5	1.8	51	3	US-09-018-635-19	Sequence 19, Appl	811	5	1.8	78	4	US-08-675-499A-28	Sequence 28, Appl
739	5	1.8	51	4	US-09-912-962-16	Sequence 16, Appl	812	5	1.8	79	4	US-09-732-210-1103	Sequence 1103, Ap
740	5	1.8	51	4	US-09-912-962-17	Sequence 17, Appl	813	5	1.8	79	4	US-09-107-532A-4870	Sequence 4870, Ap
741	5	1.8	53	1	US-09-912-962-18	Sequence 18, Appl	814	5	1.8	81	4	US-09-252-991A-24842	Sequence 24842, A
742	5	1.8	53	4	US-08-519-103-15	Sequence 15, Appl	815	5	1.8	81	4	US-09-352-991A-28421	Sequence 28421, A
743	5	1.8	53	3	US-08-905-223-489	Sequence 489, App	816	5	1.8	82	1	US-08-225-757B-12	Sequence 12, Appl
744	5	1.8	53	3	US-09-018-635-15	Sequence 15, Appl	817	5	1.8	82	4	US-09-257-179-114	Sequence 114, App
745	5	1.8	53	3	US-09-912-962-15	Sequence 15, Appl	818	5	1.8	82	4	US-09-325-932A-188	Sequence 188, App
746	5	1.8	55	1	US-08-199-776-4	Sequence 4, Appl1	819	5	1.8	83	2	US-08-499-676A-10	Sequence 10, Appl
747	5	1.8	55	3	US-08-663-731-4	Sequence 4, Appl1	820	5	1.8	83	2	US-08-499-676A-11	Sequence 11, Appl
748	5	1.8	55	3	US-08-879-338-4	Sequence 4, Appl1	821	5	1.8	83	2	US-08-499-676A-12	Sequence 12, Appl
749	5	1.8	55	4	US-09-293-238B-4	Sequence 4, Appl1	822	5	1.8	84	4	US-09-107-532A-6059	Sequence 6059, Ap
750	5	1.8	55	4	US-08-675-499A-8	Sequence 8, Appl1	823	5	1.8	84	4	US-09-134-001C-3585	Sequence 3585, Ap
751	5	1.8	55	4	US-08-675-499A-30	Sequence 30, Appl	824	5	1.8	85	4	US-09-205-258-637	Sequence 637, App
752	5	1.8	55	5	PCT-US95-02044-4	Sequence 4, Appl1	825	5	1.8	85	4	US-09-107-532A-6644	Sequence 6644, Ap
753	5	1.8	56	4	US-09-483-273-146	Sequence 146, App	826	5	1.8	86	1	US-08-014-153D-21	Sequence 21, Appl
754	5	1.8	57	5	PCT-US92-07813-25	Sequence 25, Appl	827	5	1.8	86	3	US-08-290-995-12	Sequence 12, Appl
755	5	1.8	58	4	US-09-227-357-186	Sequence 186, App	828	5	1.8	86	4	US-09-732-210-58	Sequence 58, Appl
756	5	1.8	58	4	US-08-971-089-10	Sequence 10, Appl	829	5	1.8	87	3	US-08-803-603-2	Sequence 2, Appl1
757	5	1.8	59	3	US-09-085-305-21	Sequence 21, Appl	830	5	1.8	87	4	US-09-227-357-544	Sequence 544, App

831	5	1.8	87	4	US-09-107-532A-5505	Sequence 5505, Ap	904	5	1.8	107	4	US-09-328-352-7927	Sequence 7927, Ap
832	5	1.8	88	2	US-08-576-626A-43	Sequence 43, Appl	905	5	1.8	108	2	US-08-483-695-42	Sequence 42, Appl
833	5	1.8	88	2	US-08-558-207A-504	Sequence 504, App	906	5	1.8	108	2	US-08-468-819-74	Sequence 74, Appl
834	5	1.8	88	4	US-09-222-938A-1	Sequence 1, Appl1	907	5	1.8	108	2	US-08-468-819-76	Sequence 76, Appl
835	5	1.8	88	4	US-09-222-938A-1	Sequence 1, Appl1	908	5	1.8	108	2	US-07-865-285-42	Sequence 42, Appl
836	5	1.8	89	4	US-09-252-991A-22243	Sequence 22243, A	909	5	1.8	108	2	US-08-487-231-42	Sequence 42, Appl
837	5	1.8	90	3	US-09-085-505-24	Sequence 24, Appl	910	5	1.8	108	2	US-08-602-725-29	Sequence 29, Appl
838	5	1.8	90	3	US-09-334-601-23	Sequence 23, Appl	911	5	1.8	108	3	US-09-201-912-42	Sequence 42, Appl
839	5	1.8	90	3	US-09-334-601-24	Sequence 24, Appl	912	5	1.8	108	4	US-09-134-001C-4556	Sequence 4556, Ap
840	5	1.8	90	3	US-09-334-601-25	Sequence 25, Appl	913	5	1.8	108	4	US-09-213-383-74	Sequence 74, Appl
841	5	1.8	90	4	US-09-198-452A-566	Sequence 566, App	914	5	1.8	108	4	US-09-213-383-76	Sequence 76, Appl
842	5	1.8	90	4	US-09-107-532A-3676	Sequence 3676, Ap	915	5	1.8	108	4	US-09-585-1738-30	Sequence 30, Appl
843	5	1.8	90	4	US-09-107-532A-3686	Sequence 3686, Ap	916	5	1.8	109	4	US-09-252-991A-28451	Sequence 28451, A
844	5	1.8	90	4	US-09-107-532A-7206	Sequence 7206, Ap	917	5	1.8	110	1	US-08-111-939-24	Sequence 24, Appl
845	5	1.8	91	4	US-08-858-207A-396	Sequence 396, App	918	5	1.8	110	1	US-08-487-748A-15	Sequence 15, Appl
846	5	1.8	91	4	US-09-390-134B-35	Sequence 35, Appl	919	5	1.8	110	3	US-08-398-633-15	Sequence 15, Appl
847	5	1.8	92	2	US-09-047-125-13	Sequence 13, Appl	920	5	1.8	110	3	US-08-480-070C-15	Sequence 15, Appl
848	5	1.8	92	3	US-07-736-335E-13	Sequence 13, Appl	921	5	1.8	110	3	US-08-829-525-15	Sequence 15, Appl
849	5	1.8	92	4	US-09-252-991A-25190	Sequence 25190, A	922	5	1.8	110	3	US-08-609-583A-15	Sequence 15, Appl
850	5	1.8	93	4	US-09-134-001C-4887	Sequence 4887, Ap	923	5	1.8	110	3	US-08-937-399-15	Sequence 15, Appl
851	5	1.8	93	4	US-09-615-192A-346	Sequence 346, App	924	5	1.8	110	4	US-09-387-919A-5	Sequence 5, Appl1
852	5	1.8	93	6	5169835-26	Patent No. 5169835	925	5	1.8	110	4	US-09-387-919A-6	Sequence 6, Appl1
853	5	1.8	94	2	US-08-814-030-1	Sequence 1, Appl1	926	5	1.8	110	4	US-09-310-367-15	Sequence 15, Appl
854	5	1.8	95	1	US-07-987-272A-9	Sequence 9, Appl1	927	5	1.8	110	4	US-09-032-337-15	Sequence 15, Appl
855	5	1.8	95	2	US-09-047-125-34	Sequence 34, Appl	928	5	1.8	110	4	US-09-198-452A-943	Sequence 943, App
856	5	1.8	95	3	US-07-736-335E-34	Sequence 34, Appl	929	5	1.8	110	4	US-09-464-231-15	Sequence 15, Appl
857	5	1.8	95	4	US-09-134-001C-4932	Sequence 4932, Ap	930	5	1.8	111	4	US-09-252-991A-17096	Sequence 17096, A
858	5	1.8	96	4	US-09-702-705-1831	Sequence 1831, Ap	931	5	1.8	111	4	US-09-252-991A-21167	Sequence 21167, A
859	5	1.8	96	4	US-09-736-457-1831	Sequence 1831, Ap	932	5	1.8	111	4	US-09-252-991A-26531	Sequence 26531, A
860	5	1.8	96	4	US-09-252-991A-23481	Sequence 23481, A	933	5	1.8	111	4	US-09-107-532A-5707	Sequence 5707, Ap
861	5	1.8	96	4	US-09-403-943B-24	Sequence 24, Appl	934	5	1.8	112	1	US-07-800-364B-15	Sequence 15, Appl
862	5	1.8	97	1	US-08-602-010A-22	Sequence 22, Appl	935	5	1.8	112	4	US-09-134-001C-2968	Sequence 2968, Ap
863	5	1.8	97	1	US-08-680-726A-22	Sequence 22, Appl	936	5	1.8	112	5	US-09-205-258-861	Sequence 861, App
864	5	1.8	97	3	US-09-092-409-22	Sequence 22, Appl	937	5	1.8	112	5	PCT-US91-03388-13	Sequence 13, App
865	5	1.8	97	4	US-09-422-936-73	Sequence 73, Appl	938	5	1.8	112	6	5494663-7	Patent No. 5494663
866	5	1.8	97	4	US-09-422-936-83	Sequence 83, Appl	939	5	1.8	113	4	US-09-134-001C-2839	Sequence 2839, Ap
867	5	1.8	98	3	US-08-851-843A-12	Sequence 12, Appl	940	5	1.8	113	4	US-09-205-258-317	Sequence 317, App
868	5	1.8	98	3	US-08-854-050-12	Sequence 12, Appl	941	5	1.8	113	4	US-09-107-532A-4426	Sequence 4426, Ap
869	5	1.8	98	4	US-09-430-323-12	Sequence 12, Appl	942	5	1.8	114	4	US-08-936-165A-271	Sequence 271, App
870	5	1.8	99	1	US-08-014-153D-46	Sequence 46, Appl	943	5	1.8	115	4	US-09-266-965-6	Sequence 6, Appl1
871	5	1.8	99	1	US-08-974-549A-194	Sequence 194, App	944	5	1.8	115	5	PCT-US95-13658-2	Sequence 2, Appl1
872	5	1.8	101	4	US-09-107-532A-5792	Sequence 5792, Ap	945	5	1.8	115	5	US-09-107-532A-6039	Sequence 6039, Ap
873	5	1.8	102	1	US-08-282-951-2	Sequence 2, Appl1	946	5	1.8	116	2	US-08-960-022-16	Sequence 16, Appl
874	5	1.8	102	3	US-08-478-097A-17	Sequence 17, Appl	947	5	1.8	116	3	US-09-018-635-25	Sequence 25, Appl
875	5	1.8	102	3	US-08-931-858E-162	Sequence 162, App	948	5	1.8	116	4	US-09-328-352-5926	Sequence 5926, Ap
876	5	1.8	102	3	US-08-981-739-162	Sequence 162, App	949	5	1.8	116	4	US-09-107-532A-5534	Sequence 5534, Ap
877	5	1.8	102	4	US-09-128-026-162	Sequence 162, App	950	5	1.8	116	4	US-09-912-962-25	Sequence 25, Appl
878	5	1.8	103	4	US-09-496-398-17	Sequence 17, Appl	951	5	1.8	117	4	US-09-615-192A-400	Sequence 400, App
879	5	1.8	103	4	US-08-180-371-12	Sequence 12, Appl	952	5	1.8	118	3	US-09-085-305-2	Sequence 2, Appl1
880	5	1.8	103	4	US-09-482-273-158	Sequence 158, App	953	5	1.8	118	4	US-09-252-991A-16797	Sequence 16797, A
881	5	1.8	104	4	US-08-858-207A-428	Sequence 428, App	954	5	1.8	118	4	US-09-328-352-4664	Sequence 4664, Ap
882	5	1.8	104	4	US-09-205-258-1153	Sequence 1153, Ap	955	5	1.8	118	4	US-09-328-352-8042	Sequence 8042, Ap
883	5	1.8	104	4	US-09-252-991A-23333	Sequence 23333, A	956	5	1.8	119	1	US-08-340-539A-18	Sequence 18, Appl
884	5	1.8	105	4	US-09-187-789-26	Sequence 26, Appl	957	5	1.8	119	1	US-08-340-539A-19	Sequence 19, Appl
885	5	1.8	105	4	US-09-139-600-21	Sequence 21, Appl	958	5	1.8	119	3	US-08-624-635-15	Sequence 15, Appl
886	5	1.8	105	4	US-09-732-210-781	Sequence 781, App	959	5	1.8	119	4	US-09-198-452A-1288	Sequence 1288, Ap
887	5	1.8	106	3	US-09-083-351-8	Sequence 8, Appl1	960	5	1.8	120	1	US-08-542-363-26	Sequence 26, Appl
888	5	1.8	106	3	US-09-083-351-20	Sequence 20, Appl	961	5	1.8	120	3	US-08-554-840-7	Sequence 7, Appl1
889	5	1.8	106	3	US-09-083-352-8	Sequence 8, Appl1	962	5	1.8	120	3	US-08-341-560B-2	Sequence 2, Appl1
890	5	1.8	106	3	US-09-083-352-20	Sequence 20, Appl	963	5	1.8	120	3	US-09-100-089-26	Sequence 26, Appl
891	5	1.8	106	4	US-09-252-991A-30769	Sequence 30769, A	964	5	1.8	120	4	US-09-228-986-128	Sequence 128, App
892	5	1.8	107	1	US-08-487-748A-16	Sequence 16, Appl	965	5	1.8	120	4	US-08-353-940-2	Sequence 2, Appl1
893	5	1.8	107	1	US-08-417-460-12	Sequence 12, Appl	966	5	1.8	120	4	US-08-353-940-2	Sequence 2, Appl1
894	5	1.8	107	3	US-08-398-633-16	Sequence 16, Appl	967	5	1.8	120	4	US-08-670-827-26	Sequence 26, Appl
895	5	1.8	107	3	US-08-480-070C-16	Sequence 16, Appl	968	5	1.8	120	4	US-09-332-595-7	Sequence 7, Appl1
896	5	1.8	107	3	US-08-829-525-16	Sequence 16, Appl	969	5	1.8	120	5	PCT-US93-03895-2	Sequence 2, Appl1
897	5	1.8	107	3	US-08-609-583A-16	Sequence 16, Appl	970	5	1.8	121	1	US-08-190-029A-12	Sequence 12, Appl
898	5	1.8	107	3	US-09-102-528-23	Sequence 23, Appl	971	5	1.8	121	2	US-08-462-695-12	Sequence 12, Appl
899	5	1.8	107	3	US-09-102-528-27	Sequence 27, Appl	972	5	1.8	122	2	US-08-820-170A-1	Sequence 1, Appl1
900	5	1.8	107	3	US-08-937-399-16	Sequence 16, Appl	973	5	1.8	122	2	US-08-477-451-44	Sequence 44, Appl
901	5	1.8	107	4	US-09-310-367-16	Sequence 16, Appl	974	5	1.8	122	3	US-09-055-699-1	Sequence 1, Appl1
902	5	1.8	107	4	US-09-032-337-16	Sequence 16, Appl	975	5	1.8	122	3	US-08-480-173A-56	Sequence 56, Appl
903	5	1.8	107	4	US-09-464-231-16	Sequence 16, Appl	976	5	1.8	122	3	US-08-484-408A-56	Sequence 56, Appl



```
977 5 1.8 122 3 US-09-273-565-1 Sequence 1, Appl1
978 5 1.8 122 4 US-09-565-538-1 Sequence 1, Appl1
979 5 1.8 122 4 US-09-661-468-1 Sequence 1, Appl1
980 5 1.8 122 4 US-09-976-165-1 Sequence 1, Appl1
981 5 1.8 123 1 US-07-869-912-2 Sequence 2, Appl1
982 5 1.8 123 2 US-08-853-659A-50 Sequence 50, Appl1
983 5 1.8 123 2 US-08-588-258B-1 Sequence 1, Appl1
984 5 1.8 123 3 US-08-460-505-1 Sequence 1, Appl1
985 5 1.8 123 4 US-09-199-637A-152 Sequence 152, App
986 5 1.8 123 4 US-09-328-352-4586 Sequence 4586, Ap
987 5 1.8 123 5 PCT-US96-08295-1 Sequence 1, Appl1
988 5 1.8 124 6 5466783-22 Patent No. 5466783
989 5 1.8 125 1 US-07-624-742-3 Sequence 6397, Ap
990 5 1.8 125 1 US-07-624-742-3 Sequence 3, Appl1
991 5 1.8 125 3 US-08-468-819-2 Sequence 2, Appl1
992 5 1.8 125 3 US-08-480-173A-50 Sequence 50, Appl1
993 5 1.8 125 3 US-08-480-173A-54 Sequence 54, Appl1
994 5 1.8 125 3 US-08-484-408A-50 Sequence 50, Appl1
995 5 1.8 125 3 US-08-484-408A-54 Sequence 54, Appl1
996 5 1.8 125 4 US-09-213-383-2 Sequence 2, Appl1
997 5 1.8 125 4 US-09-107-532A-5637 Sequence 5637, Ap
998 5 1.8 126 4 US-09-702-705-791 Sequence 791, App
999 5 1.8 126 4 US-09-736-457-791 Sequence 791, App
1000 5 1.8 126 6 5268359-5 Patent No. 5268359
```

## ALIGNMENTS

```
RESULT 1
US-09-610-185C-4
; Sequence 4, Application US/09610185C
; Patent No. 6573432
; GENERAL INFORMATION:
; APPLICANT: Borevitz, Justin
; APPLICANT: Xia, Yij1
; APPLICANT: Dixon, Richard A.
; APPLICANT: Lamb, Christopher J.
; TITLE OF INVENTION: REGULATION OF ANTHOCYANIN PIGMENT
; TITLE OF INVENTION: PRODUCTION
; FILE REFERENCE: SALKINS.003C1
; CURRENT APPLICATION NUMBER: US/09/610,185C
; CURRENT FILING DATE: 2000-07-05
; PRIOR APPLICATION NUMBER: US 09/603,244
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(249)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-610-185C-4

Query Match 11.7%; Score 32; DB 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 3.7e-23;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 2
US-09-610-185C-2
; Sequence 2, Application US/09610185C
; Patent No. 6573432
; GENERAL INFORMATION:
; APPLICANT: Borevitz, Justin
; APPLICANT: Xia, Yij1
```

```
APPLICANT: Dixon, Richard A.
APPLICANT: Lamb, Christopher J.
TITLE OF INVENTION: REGULATION OF ANTHOCYANIN PIGMENT
TITLE OF INVENTION: PRODUCTION
FILE REFERENCE: SALKINS.003C1
CURRENT APPLICATION NUMBER: US/09/610,185C
CURRENT FILING DATE: 2000-07-05
PRIOR APPLICATION NUMBER: US 09/603,244
PRIOR FILING DATE: 2000-06-23
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
US-09-610-185C-2

Query Match 9.9%; Score 27; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.5e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 3
US-08-472-934-10
; Sequence 10, Application US/08472934
; Patent No. 5753446
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
; TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,934
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,421
; FILING DATE: 15-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/354,516
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 5,405,941
; FILING DATE: 15-Apr-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,460
; FILING DATE: 14-Oct-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11690
; FILING DATE: 14-Oct-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04178
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. Deconti, Jr. Esq.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: CPI-004DVCPC2
```



TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1247 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-472-934-10

Query Match 2.9%; Score 8; DB 1; Length 1247;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LILRLHKL 83  
Db 183 LILRLHKL 190

RESULT 4  
US-08-323-460A-10  
Sequence 10, Application US/08323460A  
Patent No. 5854043  
GENERAL INFORMATION:  
APPLICANT: JOHNSON, GARY L.  
TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL  
TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SHERIDAN ROSS & MCINTOSH  
STREET: 1700 LINCOLN STREET, SUITE 3500  
CITY: DENVER  
STATE: CO  
COUNTRY: USA  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/323,460A  
FILING DATE: 14-OCT-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,254  
FILING DATE: 14-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: KOVARIK, JOSEPH E.  
REGISTRATION NUMBER: 33,005  
REFERENCE/DOCKET NUMBER: 2879-1-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303/863-9700  
TELEFAX: 303/863-0223  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1247 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-323-460A-10

Query Match 2.9%; Score 8; DB 2; Length 1247;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LILRLHKL 83  
Db 183 LILRLHKL 190

RESULT 5

US-08-461-146C-10  
Sequence 10, Application US/08461146C  
Patent No. 5981265  
GENERAL INFORMATION:  
APPLICANT: JOHNSON, GARY L.  
TITLE OF INVENTION: METHODS FOR REGULATING MEKK PROTEIN ACTIVITY  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Lahive and Cockfield  
STREET: 60 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,146C  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/354,516  
FILING DATE: 21-FEB-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,254  
FILING DATE: 15-APR-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/323,460  
FILING DATE: 14-OCT-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/11690  
FILING DATE: 14-OCT-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/04178  
FILING DATE: 15-APR-1994

ATTORNEY/AGENT INFORMATION:  
NAME: KARA, Catherine J.  
REGISTRATION NUMBER: P41,106  
REFERENCE/DOCKET NUMBER: CPI-004CN3

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1247 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-461-146C-10

Query Match 2.9%; Score 8; DB 2; Length 1247;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LILRLHKL 83  
Db 183 LILRLHKL 190

RESULT 6  
US-08-461-145C-10  
Sequence 10, Application US/08461145C  
Patent No. 6074861  
GENERAL INFORMATION:  
APPLICANT: JOHNSON, GARY L.  
TITLE OF INVENTION: NOVEL MEKK PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Lahive and Cockfield, LLP  
STREET: 60 State Street

CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,145C  
FILING DATE: 5-JUNE-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,254  
FILING DATE: 11-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/323,460  
FILING DATE: 14-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/11690  
FILING DATE: 14-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/04178  
FILING DATE: 15-APR-1994  
APPLICATION NUMBER: US 08/354,516  
FILING DATE: 21-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Kara, Catherine J.  
REGISTRATION NUMBER: P-41,106  
REFERENCE/DOCKET NUMBER: CPI-004CN1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1247 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-461-145C-10

Query Match 2.9%; Score 8; DB 3; Length 1247;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LILRLHKL 83  
Db 183 LILRLHKL 190

RESULT 7  
US-09-423-890-13  
Sequence 13, Application US/09423890  
Patent No. 6312934  
GENERAL INFORMATION:  
APPLICANT: CADUS PHARMACEUTICAL CORPORATION  
TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES  
FILE REFERENCE: CPI-085CPC  
CURRENT APPLICATION NUMBER: US/09/423,890  
FILING DATE: 2000-03-06  
PRIOR APPLICATION NUMBER: USSN 60/078,153  
PRIOR FILING DATE: 1998-03-16  
PRIOR APPLICATION NUMBER: USSN 60/099,165  
PRIOR FILING DATE: 1998-09-04  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 13  
LENGTH: 1597  
TYPE: PRT  
ORGANISM: Mus musculus

US-09-423-890-13

Query Match 2.9%; Score 8; DB 4; Length 1597;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LILRLHKL 83  
Db 533 LILRLHKL 540

RESULT 8  
US-08-628-829-14  
Sequence 14, Application US/08628829A  
Patent No. 6333170  
GENERAL INFORMATION:  
APPLICANT: Johnson, Gary L.  
TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To External  
FILE REFERENCE: CPI-004DVC3  
CURRENT APPLICATION NUMBER: US/08/628,829A  
FILING DATE: 1996-04-05  
EARLIER APPLICATION NUMBER: 08/440,421  
EARLIER FILING DATE: 1995-05-15  
EARLIER APPLICATION NUMBER: 08/323,460  
EARLIER FILING DATE: 1994-10-14  
EARLIER APPLICATION NUMBER: 08/049,254  
EARLIER FILING DATE: 1993-05-15  
EARLIER APPLICATION NUMBER: 08/410,602  
EARLIER FILING DATE: 1995-04-24  
EARLIER APPLICATION NUMBER: 08/472,934  
EARLIER FILING DATE: 1995-06-06  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 14  
LENGTH: 1597  
TYPE: PRT  
ORGANISM: Mus musculus  
US-08-628-829-14

Query Match 2.9%; Score 8; DB 4; Length 1597;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LILRLHKL 83  
Db 533 LILRLHKL 540

RESULT 9  
US-09-058-489-4  
Sequence 4, Application US/09058489  
Patent No. 6103866  
GENERAL INFORMATION:  
APPLICANT: Whitehead Institute for Biomedical Research  
APPLICANT: Lahn, Bruce  
TITLE OF INVENTION: Genes in the No. 6103866-Recombining Region of  
FILE REFERENCE: W197-08PA  
CURRENT APPLICATION NUMBER: US/09/058,489  
FILING DATE: 1998-04-10  
EARLIER APPLICATION NUMBER: 60/041,877  
EARLIER FILING DATE: 1997-04-11  
NUMBER OF SEQ ID NOS: 91  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 60  
TYPE: PRT  
ORGANISM: Drosophila  
US-09-058-489-4

Query Match 2.6%; Score 7; DB 3; Length 60;  
Best Local Similarity 100.0%; Pred. No. 18;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 174 EKOTPDA 180  
Db 11 EKOTPDA 17

RESULT 10  
US-09-252-991A-20768  
; Sequence 20768, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20768  
; LENGTH: 183  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20768

Query Match 2.6%; Score 7; DB 4; Length 183;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 RLPRRTA 100  
Db 85 RLPRRTA 91

RESULT 11  
US-09-252-991A-26955  
; Sequence 26955, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 26955  
; LENGTH: 218  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-26955

Query Match 2.6%; Score 7; DB 4; Length 218;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 RCRKSCR 53  
Db 46 RCRKSCR 52

RESULT 12  
US-08-722-626B-2  
; Sequence 2, Application US/08722626B  
; Patent No. 5939601

GENERAL INFORMATION:  
; APPLICANT: Yang, Yinnong  
; APPLICANT: Kleesig, Daniel, F.  
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ENHANCED  
; TITLE OF INVENTION: DISEASE RESISTANCE IN PLANTS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman  
; STREET: 1601 Market Street Suite 720  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2307  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/722,626B  
; FILING DATE: 27-SEP-1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pat Hagan  
; REGISTRATION NUMBER: 27,643  
; REFERENCE/DOCKET NUMBER: 97-0010  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215 563-4100  
; TELEFAX: 215 563-4044  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 278 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: Internal  
; ORIGINAL SOURCE:  
US-08-722-626B-2

Query Match 2.6%; Score 7; DB 2; Length 278;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 KSCRLRW 56  
Db 51 KSCRLRW 57

RESULT 13  
US-09-328-352-6229  
; Sequence 6229, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; PRIOR FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6229  
; LENGTH: 289  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6229

Query Match 2.6%; Score 7; DB 4; Length 289;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 LHHEIS 223  
Db 258 LHHEIS 264

RESULT 14  
US-09-495-406-15  
; Sequence 15, Application US/09495406  
; Patent No. 6503744  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Makarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000110US  
; CURRENT APPLICATION NUMBER: US/09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 301  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
US-09-495-406-15

Query Match 2.6%; Score 7; DB 4; Length 301;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EODEVDL 76  
Db 109 EODEVDL 115

RESULT 15  
US-08-671-320-11  
; Sequence 11, Application US/08671320  
; Patent No. 5840558  
; GENERAL INFORMATION:  
; APPLICANT: VIERLING JR, RICHARD A  
; TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN  
; TITLE OF INVENTION: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: 1125 SO. 103RD STREET  
; STREET: SUITE 330  
; CITY: OMAHA  
; STATE: NE  
; COUNTRY: US  
; ZIP: 68124-1076  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/671,320  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JONDLE, ROBERT J.  
; REGISTRATION NUMBER: 33,915  
; REFERENCE/DOCKET NUMBER: 1227-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 402-398-9000  
; TELEFAX: 402-398-9005

; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 324 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-671-320-11

Query Match 2.6%; Score 7; DB 2; Length 324;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 KINTTKI 120  
Db 233 KINTTKI 239

Search completed: January 29, 2004, 20:21:58  
Job time : 32 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 20:21:13 ; Search time 33 Seconds  
(without alignments)  
1725.570 Million cell updates/sec

Title: US-10-033-190-2

Perfect score: 274

Sequence: 1 NMSTMSLSGVKRGKSWTDEE.....OOGVNDPFSADIDMLND 274

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 789580 seqs, 207824079 residues

Word size : 0

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	274	100.0	274	US-10-033-190-2	Sequence 2, Appl1
2	32	11.7	209	US-09-934-455-480	Sequence 480, App
3	32	11.7	246	US-09-934-455-482	Sequence 482, App
4	32	11.7	249	US-09-934-455-14	Sequence 14, Appl
5	32	11.7	249	US-10-295-403-44	Sequence 44, Appl
6	27	9.9	212	US-09-934-455-438	Sequence 438, App
7	17	6.2	120	US-09-443-704-6	Sequence 6, Appl1
8	17	6.2	120	US-10-008-118A-6	Sequence 390, App
9	14	5.1	203	US-09-934-455-390	Sequence 29, Appl
10	14	5.1	203	US-10-407-920-29	Sequence 70, Appl
11	14	5.1	219	US-09-934-455-70	Sequence 22, Appl
12	13	4.7	115	US-10-021-811-22	Sequence 60, Appl
13	13	4.7	120	US-10-021-811-6	Sequence 18, Appl
14	13	4.7	131	US-10-021-811-60	
15	13	4.7	145	US-10-021-811-18	

16	4.7	146	15	US-10-021-811-20	Sequence 20, Appl
17	4.7	157	15	US-10-021-811-2	Sequence 2, Appl1
18	4.7	196	15	US-10-021-811-44	Sequence 44, Appl
19	4.7	204	15	US-10-021-811-32	Sequence 32, Appl
20	4.7	204	15	US-10-021-811-38	Sequence 38, Appl
21	4.7	206	12	US-10-225-068-2	Sequence 2, Appl1
22	4.7	206	12	US-10-302-267-130	Sequence 130, App
23	4.7	206	15	US-10-021-811-34	Sequence 34, Appl
24	4.7	206	15	US-10-021-811-16	Sequence 36, Appl
25	4.7	208	15	US-10-021-811-30	Sequence 30, Appl
26	4.7	217	15	US-10-021-811-63	Sequence 63, Appl
27	4.7	226	16	US-10-278-536-88	Sequence 88, Appl
28	4.7	226	16	US-10-278-536-86	Sequence 86, Appl
29	4.7	235	15	US-10-021-811-10	Sequence 40, Appl
30	4.7	273	11	US-09-533-029-40	Sequence 40, Appl
31	4.7	297	12	US-10-302-267-146	Sequence 146, App
32	4.7	297	12	US-10-278-173-56	Sequence 56, Appl
33	4.7	307	11	US-09-934-455-208	Sequence 208, App
34	4.7	307	12	US-10-225-068-156	Sequence 156, App
35	4.7	309	16	US-10-278-536-120	Sequence 120, App
36	4.7	322	15	US-10-021-811-46	Sequence 46, Appl
37	4.7	322	11	US-09-934-455-484	Sequence 484, App
38	4.7	329	12	US-10-302-267-144	Sequence 144, App
39	4.7	333	15	US-10-278-173-146	Sequence 146, App
40	4.7	351	15	US-10-021-811-110	Sequence 110, App
41	4.7	374	12	US-10-342-224-112	Sequence 112, App
42	4.7	376	15	US-10-278-173-142	Sequence 142, App
43	4.7	420	9	US-09-443-704-50	Sequence 50, Appl
44	4.4	120	14	US-10-008-118A-50	Sequence 50, Appl
45	4.4	149	9	US-09-443-704-30	Sequence 30, Appl
46	4.4	149	14	US-10-008-118A-30	Sequence 30, Appl
47	4.4	246	9	US-09-443-704-32	Sequence 32, Appl
48	4.4	246	14	US-10-008-118A-32	Sequence 32, Appl
49	4.4	249	15	US-10-286-264-148	Sequence 148, App
50	4.4	253	9	US-09-443-704-28	Sequence 28, Appl
51	4.4	253	14	US-10-008-118A-28	Sequence 28, Appl
52	4.4	268	9	US-09-443-704-34	Sequence 34, Appl
53	4.4	268	14	US-10-008-118A-34	Sequence 34, Appl
54	4.4	273	9	US-09-443-704-20	Sequence 20, Appl
55	4.4	273	14	US-10-008-118A-20	Sequence 20, Appl
56	4.4	281	15	US-10-295-403-46	Sequence 46, Appl
57	4.4	282	12	US-10-225-068-10	Sequence 10, Appl
58	4.4	295	9	US-09-443-704-24	Sequence 24, Appl
59	4.4	295	14	US-10-008-118A-24	Sequence 24, Appl
60	4.0	192	15	US-10-021-811-50	Sequence 50, Appl
61	4.0	229	15	US-10-021-811-42	Sequence 42, Appl
62	4.0	286	12	US-10-302-267-152	Sequence 152, App
63	4.0	312	15	US-10-021-811-48	Sequence 48, Appl
64	4.0	314	16	US-10-278-536-216	Sequence 216, App
65	3.6	103	9	US-09-443-704-44	Sequence 44, Appl
66	3.6	103	14	US-10-008-118A-44	Sequence 44, Appl
67	3.6	342	11	US-09-934-455-252	Sequence 252, App
68	3.6	363	12	US-10-093-837-2	Sequence 2, Appl1
69	3.6	371	11	US-09-533-029-86	Sequence 86, Appl
70	3.6	371	11	US-09-934-455-66	Sequence 66, Appl
71	3.6	371	16	US-10-278-536-40	Sequence 40, Appl
72	3.6	43	12	US-10-407-920-26	Sequence 26, Appl
73	2.9	74	12	US-10-407-920-16	Sequence 16, Appl
74	2.9	75	12	US-10-407-920-18	Sequence 18, Appl
75	2.9	159	12	US-10-097-111-322	Sequence 32, App
76	2.9	281	12	US-10-264-237-1988	Sequence 1988, App
77	2.9	289	10	US-09-738-626-4917	Sequence 4917, App
78	2.9	295	15	US-10-278-173-152	Sequence 152, App
79	2.9	1597	14	US-10-000-864-13	Sequence 13, Appl
80	2.6	54	12	US-10-029-386-58317	Sequence 28317, A
81	2.6	54	15	US-10-021-811-56	Sequence 56, Appl
82	2.6	62	9	US-09-050-010-10	Sequence 10, Appl
83	2.6	62	15	US-10-021-811-16	Sequence 16, Appl
84	2.6	71	15	US-10-021-811-12	Sequence 12, Appl
85	2.6	90	15	US-10-021-811-52	Sequence 52, Appl
86	2.6	94	15	US-10-021-811-28	Sequence 28, Appl
87	2.6	110	9	US-09-443-704-26	Sequence 26, Appl
88	2.6				

89	7	2.6	110	14	US-10-008-118A-26	Sequence 26, App1	162	7	2.6	305	14	US-10-008-118A-8	Sequence 8, App1
90	7	2.6	115	9	US-09-443-704-36	Sequence 36, App1	163	7	2.6	307	15	US-10-278-173-18	Sequence 148, App1
91	7	2.6	115	14	US-10-008-118A-36	Sequence 36, App1	164	7	2.6	307	16	US-10-278-536-92	Sequence 92, App1
92	7	2.6	122	11	US-09-533-029-112	Sequence 112, App1	165	7	2.6	309	16	US-10-278-536-208	Sequence 208, App1
93	7	2.6	122	15	US-10-021-811-62	Sequence 62, App1	166	7	2.6	314	12	US-10-302-267-150	Sequence 150, App1
94	7	2.6	122	16	US-10-278-536-214	Sequence 214, App1	167	7	2.6	314	15	US-10-278-173-58	Sequence 58, App1
95	7	2.6	126	15	US-10-021-811-24	Sequence 24, App1	168	7	2.6	314	15	US-10-278-536-66	Sequence 66, App1
96	7	2.6	128	9	US-09-443-704-2	Sequence 2, App1	169	7	2.6	317	11	US-09-934-455-156	Sequence 156, App1
97	7	2.6	128	14	US-10-008-118A-2	Sequence 2, App1	170	7	2.6	319	11	US-09-934-455-366	Sequence 366, App1
98	7	2.6	136	15	US-10-278-173-134	Sequence 134, App1	171	7	2.6	320	15	US-10-278-173-68	Sequence 68, App1
99	7	2.6	139	9	US-09-443-704-38	Sequence 38, App1	172	7	2.6	320	15	US-10-278-536-118	Sequence 118, App1
100	7	2.6	139	14	US-10-008-118A-38	Sequence 38, App1	173	7	2.6	320	15	US-10-021-811-26	Sequence 26, App1
101	7	2.6	144	15	US-10-083-357-755	Sequence 755, App1	174	7	2.6	323	15	US-10-286-264-84	Sequence 84, App1
102	7	2.6	179	15	US-10-021-811-4	Sequence 4, App1	175	7	2.6	323	15	US-10-295-403-66	Sequence 66, App1
103	7	2.6	180	12	US-10-320-797-3023	Sequence 3023, App1	176	7	2.6	324	16	US-10-278-536-68	Sequence 68, App1
104	7	2.6	188	9	US-09-443-704-22	Sequence 22, App1	177	7	2.6	326	11	US-09-934-455-364	Sequence 364, App1
105	7	2.6	188	14	US-10-008-118A-22	Sequence 22, App1	178	7	2.6	327	11	US-09-533-029-68	Sequence 68, App1
106	7	2.6	191	14	US-10-286-264-126	Sequence 126, App1	179	7	2.6	327	11	US-09-533-029-68	Sequence 68, App1
107	7	2.6	198	15	US-10-021-811-58	Sequence 58, App1	180	7	2.6	327	16	US-10-278-536-14	Sequence 14, App1
108	7	2.6	217	9	US-09-443-704-14	Sequence 14, App1	181	7	2.6	332	12	US-10-225-068-6	Sequence 6, App1
109	7	2.6	217	14	US-10-008-118A-14	Sequence 14, App1	182	7	2.6	332	12	US-10-021-811-54	Sequence 54, App1
110	7	2.6	221	15	US-10-021-811-14	Sequence 14, App1	183	7	2.6	333	15	US-10-295-403-42	Sequence 42, App1
111	7	2.6	238	11	US-09-934-455-198	Sequence 198, App1	184	7	2.6	335	15	US-10-295-403-58	Sequence 58, App1
112	7	2.6	238	12	US-10-302-267-140	Sequence 140, App1	185	7	2.6	336	11	US-09-533-029-58	Sequence 58, App1
113	7	2.6	239	12	US-10-225-068-30	Sequence 30, App1	186	7	2.6	336	11	US-10-286-264-94	Sequence 94, App1
114	7	2.6	243	11	US-09-934-455-202	Sequence 202, App1	187	7	2.6	336	15	US-10-295-403-62	Sequence 62, App1
115	7	2.6	245	15	US-10-295-403-60	Sequence 60, App1	188	7	2.6	339	15	US-10-278-173-138	Sequence 138, App1
116	7	2.6	249	11	US-09-934-455-436	Sequence 436, App1	189	7	2.6	349	16	US-10-278-536-114	Sequence 114, App1
117	7	2.6	249	15	US-10-286-264-132	Sequence 132, App1	190	7	2.6	350	16	US-09-935-943-6	Sequence 6, App1
118	7	2.6	250	11	US-09-934-455-206	Sequence 206, App1	191	7	2.6	352	11	US-09-934-455-118	Sequence 118, App1
119	7	2.6	250	12	US-10-302-267-134	Sequence 134, App1	192	7	2.6	352	16	US-10-278-173-110	Sequence 110, App1
120	7	2.6	258	9	US-09-443-704-4	Sequence 4, App1	193	7	2.6	352	16	US-10-278-536-80	Sequence 80, App1
121	7	2.6	258	14	US-10-008-118A-4	Sequence 4, App1	194	7	2.6	359	12	US-10-225-068-32	Sequence 32, App1
122	7	2.6	264	9	US-09-443-704-16	Sequence 16, App1	195	7	2.6	362	16	US-10-278-536-80	Sequence 80, App1
123	7	2.6	264	14	US-10-008-118A-16	Sequence 16, App1	196	7	2.6	365	15	US-10-295-403-52	Sequence 52, App1
124	7	2.6	268	11	US-09-934-455-324	Sequence 324, App1	197	7	2.6	367	16	US-10-278-536-206	Sequence 206, App1
125	7	2.6	268	15	US-10-278-173-2	Sequence 2, App1	198	7	2.6	367	16	US-10-225-068-90	Sequence 90, App1
126	7	2.6	268	15	US-10-278-173-26	Sequence 26, App1	199	7	2.6	367	16	US-10-225-068-90	Sequence 90, App1
127	7	2.6	269	11	US-09-533-029-94	Sequence 94, App1	200	7	2.6	432	12	US-10-302-267-142	Sequence 142, App1
128	7	2.6	269	15	US-10-278-173-72	Sequence 72, App1	201	7	2.6	432	12	US-10-302-267-142	Sequence 142, App1
129	7	2.6	273	11	US-09-533-029-62	Sequence 62, App1	202	7	2.6	432	12	US-10-369-493-18430	Sequence 18430, App1
130	7	2.6	273	11	US-09-934-455-204	Sequence 204, App1	203	7	2.6	432	12	US-10-369-493-19176	Sequence 19176, App1
131	7	2.6	273	15	US-10-278-173-140	Sequence 140, App1	204	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
132	7	2.6	273	15	US-10-286-264-128	Sequence 128, App1	205	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
133	7	2.6	273	16	US-10-278-536-58	Sequence 58, App1	206	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
134	7	2.6	276	12	US-10-302-267-164	Sequence 164, App1	207	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
135	7	2.6	276	12	US-10-302-267-164	Sequence 164, App1	208	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
136	7	2.6	280	11	US-09-934-455-200	Sequence 200, App1	209	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
137	7	2.6	280	12	US-10-302-267-132	Sequence 132, App1	210	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
138	7	2.6	282	12	US-10-302-267-148	Sequence 148, App1	211	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
139	7	2.6	282	15	US-10-295-403-54	Sequence 54, App1	212	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
140	7	2.6	283	11	US-09-934-455-368	Sequence 368, App1	213	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
141	7	2.6	283	12	US-10-225-068-214	Sequence 214, App1	214	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
142	7	2.6	283	15	US-10-278-173-154	Sequence 154, App1	215	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
143	7	2.6	286	11	US-09-934-455-510	Sequence 510, App1	216	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
144	7	2.6	293	12	US-10-306-762-178	Sequence 178, App1	217	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
145	7	2.6	299	12	US-10-302-267-138	Sequence 138, App1	218	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
146	7	2.6	300	12	US-10-302-267-154	Sequence 154, App1	219	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
147	7	2.6	301	9	US-09-816-028A-27	Sequence 27, App1	220	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
148	7	2.6	301	12	US-10-303-161-27	Sequence 27, App1	221	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
149	7	2.6	301	12	US-10-303-118-27	Sequence 27, App1	222	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
150	7	2.6	301	12	US-10-303-128-27	Sequence 27, App1	223	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
151	7	2.6	301	12	US-10-303-134-27	Sequence 27, App1	224	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
152	7	2.6	301	12	US-10-303-162-27	Sequence 27, App1	225	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
153	7	2.6	302	9	US-09-443-704-12	Sequence 12, App1	226	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
154	7	2.6	302	9	US-09-443-704-18	Sequence 18, App1	227	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
155	7	2.6	302	14	US-10-008-118A-12	Sequence 12, App1	228	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
156	7	2.6	302	14	US-10-008-118A-18	Sequence 18, App1	229	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
157	7	2.6	304	11	US-09-934-455-64	Sequence 64, App1	230	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
158	7	2.6	304	12	US-10-225-068-188	Sequence 188, App1	231	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
159	7	2.6	305	9	US-09-443-704-8	Sequence 8, App1	232	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
160	7	2.6	305	11	US-09-533-029-32	Sequence 32, App1	233	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1
161	7	2.6	305	11	US-09-934-455-68	Sequence 68, App1	234	7	2.6	432	12	US-10-369-493-2112	Sequence 2112, App1

235	6	2.2	55	9	US-09-864-761-34815	Sequence 34815, A	308	6	2.2	250	11	US-09-308-207-17	Sequence 17, Appl
236	6	2.2	57	10	US-09-738-626-5498	Sequence 5498, Ap	309	6	2.2	250	12	US-10-277-249-56	Sequence 56, Appl
237	6	2.2	65	12	US-09-864-408A-8784	Sequence 8784, Ap	310	6	2.2	250	12	US-10-369-493-1746	Sequence 1746, Ap
238	6	2.2	66	14	US-10-016-157A-178	Sequence 178, App	311	6	2.2	250	12	US-10-108-260A-2691	Sequence 2691, Ap
239	6	2.2	68	12	US-09-864-408A-5102	Sequence 5102, Ap	312	6	2.2	259	15	US-10-304-928-18	Sequence 18, Appl
240	6	2.2	70	9	US-09-864-761-39034	Sequence 39034, A	313	6	2.2	261	12	US-10-032-201B-162	Sequence 162, Appl
241	6	2.2	70	14	US-10-016-157A-179	Sequence 179, App	314	6	2.2	261	12	US-10-369-493-6381	Sequence 6381, Ap
242	6	2.2	73	8	US-08-424-5508-510	Sequence 510, App	315	6	2.2	262	15	US-10-253-007-35	Sequence 35, Appl
243	6	2.2	73	9	US-10-029-386-33789	Sequence 33789, A	316	6	2.2	263	12	US-10-137-018-484	Sequence 484, App
244	6	2.2	78	12	US-09-764-869-769	Sequence 769, App	317	6	2.2	263	12	US-10-140-021-484	Sequence 484, App
245	6	2.2	78	12	US-10-027-577-769	Sequence 769, App	318	6	2.2	263	12	US-10-140-021-484	Sequence 484, App
246	6	2.2	78	15	US-10-091-504-769	Sequence 769, App	319	6	2.2	263	12	US-10-140-021-484	Sequence 484, App
247	6	2.2	82	10	US-09-764-864-1460	Sequence 1460, Ap	320	6	2.2	263	12	US-10-140-021-484	Sequence 484, App
248	6	2.2	82	11	US-09-764-891-4569	Sequence 4569, Ap	321	6	2.2	263	12	US-10-140-021-484	Sequence 484, App
249	6	2.2	84	11	US-09-764-891-4455	Sequence 4455, Ap	322	6	2.2	263	12	US-10-140-021-484	Sequence 484, App
250	6	2.2	84	11	US-09-764-891-4817	Sequence 4817, Ap	323	6	2.2	263	12	US-10-140-021-484	Sequence 484, App
251	6	2.2	84	15	US-10-091-572-372	Sequence 372, App	324	6	2.2	263	12	US-10-140-021-484	Sequence 484, App
252	6	2.2	86	15	US-10-264-049-2350	Sequence 2350, App	325	6	2.2	263	12	US-10-141-696-484	Sequence 484, App
253	6	2.2	87	15	US-10-156-761-10764	Sequence 10764, A	326	6	2.2	263	12	US-10-141-702-484	Sequence 484, App
254	6	2.2	91	15	US-10-106-698-5428	Sequence 5428, Ap	327	6	2.2	263	12	US-10-141-702-484	Sequence 484, App
255	6	2.2	94	15	US-10-156-761-9738	Sequence 9738, Ap	328	6	2.2	263	12	US-10-142-421-484	Sequence 484, App
256	6	2.2	102	12	US-10-187-394-18	Sequence 18, Appl	329	6	2.2	263	12	US-10-142-421-484	Sequence 484, App
257	6	2.2	103	12	US-10-104-047-3124	Sequence 3124, Ap	330	6	2.2	263	12	US-10-142-767-484	Sequence 484, App
258	6	2.2	103	12	US-10-289-762-1230	Sequence 1230, Ap	331	6	2.2	263	12	US-10-143-033-484	Sequence 484, App
259	6	2.2	110	9	US-09-864-761-34071	Sequence 34071, A	332	6	2.2	263	12	US-10-144-924-484	Sequence 484, App
260	6	2.2	123	15	US-10-029-386-29996	Sequence 29996, A	333	6	2.2	263	12	US-10-145-628-484	Sequence 484, App
261	6	2.2	123	15	US-10-017-161-1442	Sequence 45991, A	334	6	2.2	263	12	US-10-145-631-484	Sequence 484, App
262	6	2.2	127	9	US-09-864-408A-4860	Sequence 4860, Ap	335	6	2.2	263	12	US-10-145-631-484	Sequence 484, App
263	6	2.2	127	12	US-10-369-493-4334	Sequence 4334, Ap	336	6	2.2	263	12	US-10-145-746-484	Sequence 484, App
264	6	2.2	141	12	US-10-369-493-7090	Sequence 7090, Ap	337	6	2.2	263	12	US-10-145-746-484	Sequence 484, App
265	6	2.2	150	12	US-10-108-260A-4133	Sequence 4133, Ap	338	6	2.2	263	12	US-10-145-823-484	Sequence 484, App
266	6	2.2	151	12	US-10-131-487A-169	Sequence 169, App	339	6	2.2	263	12	US-10-145-823-484	Sequence 484, App
267	6	2.2	159	12	US-09-764-868-864	Sequence 864, App	340	6	2.2	263	12	US-10-145-870-484	Sequence 484, App
268	6	2.2	159	12	US-10-353-174-4	Sequence 4, Appl	341	6	2.2	263	12	US-10-145-870-484	Sequence 484, App
269	6	2.2	165	12	US-10-108-260A-2473	Sequence 2473, Ap	342	6	2.2	263	12	US-10-145-955-484	Sequence 484, App
270	6	2.2	168	10	US-09-738-626-5722	Sequence 5722, Ap	343	6	2.2	263	12	US-10-146-724-484	Sequence 484, App
271	6	2.2	169	9	US-09-956-425-2	Sequence 2, Appl	344	6	2.2	263	12	US-10-146-724-484	Sequence 484, App
272	6	2.2	172	12	US-10-292-798-1168	Sequence 1168, Ap	345	6	2.2	263	12	US-10-146-724-484	Sequence 484, App
273	6	2.2	172	12	US-10-369-493-9133	Sequence 9133, Ap	346	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
274	6	2.2	179	12	US-10-369-493-9462	Sequence 9462, Ap	347	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
275	6	2.2	182	12	US-10-017-161-1646	Sequence 1646, Ap	348	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
276	6	2.2	182	12	US-10-292-798-1314	Sequence 1314, Ap	349	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
277	6	2.2	187	12	US-09-864-408A-4876	Sequence 4876, Ap	350	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
278	6	2.2	194	10	US-09-738-626-5004	Sequence 5004, Ap	351	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
279	6	2.2	196	11	US-09-883-802-393	Sequence 393, App	352	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
280	6	2.2	199	12	US-10-289-762-1204	Sequence 1204, Ap	353	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
281	6	2.2	199	10	US-09-738-626-4571	Sequence 4571, Ap	354	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
282	6	2.2	205	15	US-10-156-761-10351	Sequence 10351, A	355	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
283	6	2.2	206	12	US-10-319-799-34	Sequence 34, Appl	356	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
284	6	2.2	211	15	US-10-189-346-50	Sequence 50, Appl	357	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
285	6	2.2	211	15	US-10-189-346-51	Sequence 51, Appl	358	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
286	6	2.2	215	12	US-10-032-585-7264	Sequence 7264, Ap	359	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
287	6	2.2	215	11	US-09-866-050A-494	Sequence 494, App	360	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
288	6	2.2	216	15	US-10-286-264-100	Sequence 100, App	361	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
289	6	2.2	216	16	US-10-378-536-156	Sequence 156, App	362	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
290	6	2.2	219	10	US-09-738-626-3534	Sequence 3534, Ap	363	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
291	6	2.2	220	10	US-09-925-300-1567	Sequence 1567, Ap	364	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
292	6	2.2	220	12	US-09-769-744A-126	Sequence 126, App	365	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
293	6	2.2	222	15	US-10-156-761-7686	Sequence 7686, Ap	366	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
294	6	2.2	222	15	US-10-369-493-17572	Sequence 17572, A	367	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
295	6	2.2	226	11	US-09-883-802-401	Sequence 401, App	368	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
296	6	2.2	226	11	US-09-883-802-401	Sequence 401, App	369	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
297	6	2.2	236	10	US-09-925-300-1191	Sequence 1191, Ap	370	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
298	6	2.2	244	12	US-10-369-493-15893	Sequence 15893, A	371	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
299	6	2.2	244	12	US-10-369-493-16266	Sequence 16266, A	372	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
300	6	2.2	248	12	US-10-316-253-161	Sequence 161, App	373	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
301	6	2.2	248	14	US-10-033-078-14	Sequence 14, Appl	374	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
302	6	2.2	249	15	US-10-369-493-15522	Sequence 15522, A	375	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
303	6	2.2	249	15	US-10-278-173-22	Sequence 22, Appl	376	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
304	6	2.2	250	10	US-09-764-864-1516	Sequence 1516, Ap	377	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
305	6	2.2					378	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
306	6	2.2					379	6	2.2	263	12	US-10-147-501-484	Sequence 484, App
307	6	2.2					380	6	2.2	263	12	US-10-147-501-484	Sequence 484, App

[illegible]



527	6	2.2	263	12	US-10-156-842-484	Sequence 484, App	600	6	2.2	263	15	US-10-066-494-41	Sequence 41, Appl
528	6	2.2	263	12	US-10-156-844-484	Sequence 484, App	601	6	2.2	263	15	US-10-137-865-484	Sequence 484, App
529	6	2.2	263	12	US-10-156-845-484	Sequence 484, App	602	6	2.2	263	15	US-10-140-477-484	Sequence 484, App
530	6	2.2	263	12	US-10-156-846-484	Sequence 484, App	603	6	2.2	263	15	US-10-142-431-484	Sequence 484, App
531	6	2.2	263	12	US-10-123-913-484	Sequence 484, App	604	6	2.2	263	15	US-10-143-114-484	Sequence 484, App
532	6	2.2	263	12	US-10-139-963-484	Sequence 484, App	605	6	2.2	263	15	US-10-140-002-484	Sequence 484, App
533	6	2.2	263	12	US-10-140-020-484	Sequence 484, App	606	6	2.2	263	15	US-10-066-269-41	Sequence 41, Appl
534	6	2.2	263	12	US-10-140-023-484	Sequence 484, App	607	6	2.2	263	15	US-10-066-211-41	Sequence 41, Appl
535	6	2.2	263	12	US-10-140-473-484	Sequence 484, App	608	6	2.2	263	15	US-10-066-193-41	Sequence 41, Appl
536	6	2.2	263	12	US-10-140-806-484	Sequence 484, App	609	6	2.2	263	15	US-10-142-419-484	Sequence 484, App
537	6	2.2	263	12	US-10-140-809-484	Sequence 484, App	610	6	2.2	263	15	US-10-123-267-484	Sequence 484, App
538	6	2.2	263	12	US-10-140-810-484	Sequence 484, App	611	6	2.2	263	15	US-10-142-423-484	Sequence 484, App
539	6	2.2	263	12	US-10-140-863-484	Sequence 484, App	612	6	2.2	263	15	US-10-121-050-484	Sequence 484, App
540	6	2.2	263	12	US-10-140-865-484	Sequence 484, App	613	6	2.2	263	15	US-10-141-755-484	Sequence 484, App
541	6	2.2	263	12	US-10-141-699-484	Sequence 484, App	614	6	2.2	263	15	US-10-213-145-4	Sequence 4, Appl
542	6	2.2	263	12	US-10-141-701-484	Sequence 484, App	615	6	2.2	263	15	US-10-143-032-484	Sequence 484, App
543	6	2.2	263	12	US-10-141-703-484	Sequence 484, App	616	6	2.2	263	15	US-10-123-108-484	Sequence 484, App
544	6	2.2	263	12	US-10-141-706-484	Sequence 484, App	617	6	2.2	263	15	US-10-123-235-484	Sequence 484, App
545	6	2.2	263	12	US-10-141-754-484	Sequence 484, App	618	6	2.2	263	15	US-10-123-261-484	Sequence 484, App
546	6	2.2	263	12	US-10-141-757-484	Sequence 484, App	619	6	2.2	263	15	US-10-140-921-484	Sequence 484, App
547	6	2.2	263	12	US-10-141-760-484	Sequence 484, App	620	6	2.2	263	15	US-10-140-928-484	Sequence 484, App
548	6	2.2	263	12	US-10-141-762-484	Sequence 484, App	621	6	2.2	263	15	US-10-121-048-484	Sequence 484, App
549	6	2.2	263	12	US-10-142-423-484	Sequence 484, App	622	6	2.2	263	15	US-10-123-292-484	Sequence 484, App
550	6	2.2	263	12	US-10-142-428-484	Sequence 484, App	623	6	2.2	263	15	US-10-123-902-484	Sequence 484, App
551	6	2.2	263	12	US-10-142-429-484	Sequence 484, App	624	6	2.2	263	15	US-10-124-819-484	Sequence 484, App
552	6	2.2	263	12	US-10-142-430-484	Sequence 484, App	625	6	2.2	263	15	US-10-124-822-484	Sequence 484, App
553	6	2.2	263	12	US-10-142-884-484	Sequence 484, App	626	6	2.2	263	15	US-10-140-925-484	Sequence 484, App
554	6	2.2	263	12	US-10-143-027-484	Sequence 484, App	627	6	2.2	263	15	US-10-160-499-484	Sequence 484, App
555	6	2.2	263	12	US-10-143-113-484	Sequence 484, App	628	6	2.2	263	15	US-10-124-824-484	Sequence 484, App
556	6	2.2	263	12	US-10-143-115-484	Sequence 484, App	629	6	2.2	263	15	US-10-127-825A-484	Sequence 484, App
557	6	2.2	263	12	US-10-144-956-484	Sequence 484, App	630	6	2.2	263	15	US-10-127-829A-484	Sequence 484, App
558	6	2.2	263	12	US-10-144-958-484	Sequence 484, App	631	6	2.2	263	15	US-10-127-835A-484	Sequence 484, App
559	6	2.2	263	12	US-10-145-632-484	Sequence 484, App	632	6	2.2	263	15	US-10-127-901A-484	Sequence 484, App
560	6	2.2	263	12	US-10-145-749-484	Sequence 484, App	633	6	2.2	263	15	US-10-128-693A-484	Sequence 484, App
561	6	2.2	263	12	US-10-145-753-484	Sequence 484, App	634	6	2.2	263	15	US-10-131-818A-484	Sequence 484, App
562	6	2.2	263	12	US-10-145-871-484	Sequence 484, App	635	6	2.2	263	15	US-10-131-819A-484	Sequence 484, App
563	6	2.2	263	12	US-10-145-878-484	Sequence 484, App	636	6	2.2	263	15	US-10-137-872A-484	Sequence 484, App
564	6	2.2	263	12	US-10-146-730-484	Sequence 484, App	637	6	2.2	263	15	US-10-147-500-484	Sequence 484, App
565	6	2.2	263	12	US-10-146-792-484	Sequence 484, App	638	6	2.2	263	15	US-10-131-822A-484	Sequence 484, App
566	6	2.2	263	12	US-10-146-794-484	Sequence 484, App	639	6	2.2	263	15	US-10-131-823A-484	Sequence 484, App
567	6	2.2	263	12	US-10-147-489-484	Sequence 484, App	640	6	2.2	263	15	US-10-131-830A-484	Sequence 484, App
568	6	2.2	263	12	US-10-147-507-484	Sequence 484, App	641	6	2.2	263	15	US-10-131-831A-484	Sequence 484, App
569	6	2.2	263	12	US-10-147-535-484	Sequence 484, App	642	6	2.2	263	15	US-10-137-872A-484	Sequence 484, App
570	6	2.2	263	12	US-10-147-537-484	Sequence 484, App	643	6	2.2	263	15	US-10-147-500-484	Sequence 484, App
571	6	2.2	263	12	US-10-152-376-484	Sequence 484, App	644	6	2.2	263	15	US-10-147-515-484	Sequence 484, App
572	6	2.2	263	12	US-10-152-381-484	Sequence 484, App	645	6	2.2	263	15	US-10-147-517-484	Sequence 484, App
573	6	2.2	263	12	US-10-152-400-484	Sequence 484, App	646	6	2.2	263	15	US-10-147-526-484	Sequence 484, App
574	6	2.2	263	12	US-10-153-585-484	Sequence 484, App	647	6	2.2	263	15	US-10-147-528-484	Sequence 484, App
575	6	2.2	263	12	US-10-156-843-484	Sequence 484, App	648	6	2.2	263	15	US-10-121-041-484	Sequence 484, App
576	6	2.2	263	12	US-10-157-780-484	Sequence 484, App	649	6	2.2	263	15	US-10-121-043-484	Sequence 484, App
577	6	2.2	263	12	US-10-157-786-484	Sequence 484, App	650	6	2.2	263	15	US-10-121-044-484	Sequence 484, App
578	6	2.2	263	12	US-10-157-800-484	Sequence 484, App	651	6	2.2	263	15	US-10-123-213-484	Sequence 484, App
579	6	2.2	263	12	US-10-157-801-484	Sequence 484, App	652	6	2.2	263	15	US-10-123-902-484	Sequence 484, App
580	6	2.2	263	12	US-10-157-802-484	Sequence 484, App	653	6	2.2	263	15	US-10-123-908-484	Sequence 484, App
581	6	2.2	263	12	US-10-158-784-484	Sequence 484, App	654	6	2.2	263	15	US-10-123-909-484	Sequence 484, App
582	6	2.2	263	12	US-10-158-789-484	Sequence 484, App	655	6	2.2	263	15	US-10-123-910-484	Sequence 484, App
583	6	2.2	263	12	US-10-158-791-484	Sequence 484, App	656	6	2.2	263	15	US-10-124-817-484	Sequence 484, App
584	6	2.2	263	12	US-10-192-011-484	Sequence 484, App	657	6	2.2	263	15	US-10-124-817-484	Sequence 484, App
585	6	2.2	263	12	US-10-147-528-484	Sequence 484, App	658	6	2.2	263	15	US-10-125-922-484	Sequence 484, App
586	6	2.2	263	12	US-10-128-692A-484	Sequence 484, App	659	6	2.2	263	15	US-10-125-922-484	Sequence 484, App
587	6	2.2	263	12	US-10-140-927-484	Sequence 484, App	660	6	2.2	263	15	US-10-140-860-484	Sequence 484, App
588	6	2.2	263	14	US-10-101-927-484	Sequence 2, Appl	661	6	2.2	263	15	US-10-142-417-484	Sequence 484, App
589	6	2.2	263	14	US-10-066-500-41	Sequence 41, Appl	662	6	2.2	263	15	US-10-147-513-484	Sequence 484, App
590	6	2.2	263	14	US-10-053-107-4	Sequence 4, Appl	663	6	2.2	263	15	US-10-157-782-484	Sequence 484, App
591	6	2.2	263	15	US-10-028-072-484	Sequence 484, App	664	6	2.2	263	15	US-10-152-395-484	Sequence 484, App
592	6	2.2	263	15	US-10-121-049-484	Sequence 484, App	665	6	2.2	263	15	US-10-123-199-4	Sequence 4, Appl
593	6	2.2	263	15	US-10-123-904-484	Sequence 484, App	666	6	2.2	263	15	US-10-125-926A-484	Sequence 484, App
594	6	2.2	263	15	US-10-140-470-484	Sequence 484, App	667	6	2.2	263	15	US-10-125-930A-484	Sequence 484, App
595	6	2.2	263	15	US-10-175-746-484	Sequence 484, App	668	6	2.2	263	15	US-10-127-831A-484	Sequence 484, App
596	6	2.2	263	15	US-10-176-918-484	Sequence 484, App	669	6	2.2	263	15	US-10-127-833A-484	Sequence 484, App
597	6	2.2	263	15	US-10-176-921-484	Sequence 484, App	670	6	2.2	263	15	US-10-127-833B-484	Sequence 484, App
598	6	2.2	263	15	US-10-002-796-41	Sequence 41, Appl	671	6	2.2	263	15	US-10-127-842A-484	Sequence 484, App
599	6	2.2	263	15	US-10-066-273-41	Sequence 41, Appl	672	6	2.2	263	15	US-10-127-843A-484	Sequence 484, App

673	6	2.2	263	15	US-10-127-845A-484	Sequence 484, App	746	6	2.2	263	15	US-10-223-083-182	Sequence 182, App
674	6	2.2	263	15	US-10-127-846A-484	Sequence 484, App	747	6	2.2	263	15	US-10-175-559-484	Sequence 484, App
675	6	2.2	263	15	US-10-127-848A-484	Sequence 484, App	748	6	2.2	263	15	US-10-223-089-182	Sequence 182, App
676	6	2.2	263	15	US-10-127-849A-484	Sequence 484, App	749	6	2.2	263	16	US-10-137-866-484	Sequence 484, App
677	6	2.2	263	15	US-10-127-850A-484	Sequence 484, App	750	6	2.2	263	16	US-10-146-726-484	Sequence 484, App
678	6	2.2	263	15	US-10-127-851A-484	Sequence 484, App	751	6	2.2	263	16	US-10-146-727-484	Sequence 484, App
679	6	2.2	263	15	US-10-128-684A-484	Sequence 484, App	752	6	2.2	263	16	US-10-146-788-484	Sequence 484, App
680	6	2.2	263	15	US-10-128-686A-484	Sequence 484, App	753	6	2.2	263	16	US-10-152-380-484	Sequence 484, App
681	6	2.2	263	15	US-10-128-690A-484	Sequence 484, App	754	6	2.2	263	16	US-10-153-934-484	Sequence 484, App
682	6	2.2	263	15	US-10-128-691A-484	Sequence 484, App	755	6	2.2	271	11	US-09-308-207-33	Sequence 33, App1
683	6	2.2	263	15	US-10-131-819A-484	Sequence 484, App	756	6	2.2	271	12	US-10-369-493-22102	Sequence 22102, A
684	6	2.2	263	15	US-10-131-829A-484	Sequence 484, App	757	6	2.2	274	12	US-10-369-493-13320	Sequence 13320, A
685	6	2.2	263	15	US-10-131-836A-484	Sequence 484, App	758	6	2.2	274	12	US-10-264-049-2732	Sequence 2732, App
686	6	2.2	263	15	US-10-146-729-484	Sequence 484, App	759	6	2.2	276	12	US-10-156-761-9562	Sequence 9562, App
687	6	2.2	263	15	US-10-146-791-484	Sequence 484, App	760	6	2.2	277	15	US-10-017-162-1576	Sequence 1576, App
688	6	2.2	263	15	US-10-147-484-484	Sequence 484, App	761	6	2.2	278	9	US-09-764-853-559	Sequence 559, App
689	6	2.2	263	15	US-10-147-508-484	Sequence 484, App	762	6	2.2	284	12	US-10-080-608A-43	Sequence 43, App1
690	6	2.2	263	15	US-10-147-512-484	Sequence 484, App	763	6	2.2	284	12	US-10-370-685-132	Sequence 132, App
691	6	2.2	263	15	US-10-175-735-484	Sequence 484, App	764	6	2.2	285	15	US-10-156-761-9562	Sequence 9562, App
692	6	2.2	263	15	US-10-121-040-484	Sequence 484, App	765	6	2.2	287	15	US-10-081-816-22	Sequence 22, App1
693	6	2.2	263	15	US-10-121-056-484	Sequence 484, App	766	6	2.2	288	12	US-10-369-493-16440	Sequence 16440, A
694	6	2.2	263	15	US-10-121-061-484	Sequence 484, App	767	6	2.2	290	12	US-10-369-493-13439	Sequence 13439, A
695	6	2.2	263	15	US-10-123-235-484	Sequence 484, App	768	6	2.2	294	9	US-09-826-508-4	Sequence 4, App1
696	6	2.2	263	15	US-10-124-818-484	Sequence 484, App	769	6	2.2	305	12	US-10-317-832-105	Sequence 105, App
697	6	2.2	263	15	US-10-137-868-484	Sequence 484, App	770	6	2.2	305	12	US-10-317-832-106	Sequence 106, App
698	6	2.2	263	15	US-10-147-492-484	Sequence 484, App	771	6	2.2	306	16	US-10-138-098-50	Sequence 50, App1
699	6	2.2	263	15	US-10-158-782-484	Sequence 484, App	772	6	2.2	308	12	US-10-306-762-22	Sequence 22, App1
700	6	2.2	263	15	US-10-123-905-484	Sequence 484, App	773	6	2.2	313	9	US-09-989-722-414	Sequence 414, App
701	6	2.2	263	15	US-10-123-907-484	Sequence 484, App	774	6	2.2	313	9	US-09-989-723-414	Sequence 414, App
702	6	2.2	263	15	US-10-124-815-484	Sequence 484, App	775	6	2.2	313	9	US-09-989-727-414	Sequence 414, App
703	6	2.2	263	15	US-10-125-921A-484	Sequence 484, App	776	6	2.2	313	9	US-09-989-727-414	Sequence 414, App
704	6	2.2	263	15	US-10-125-928A-484	Sequence 484, App	777	6	2.2	313	10	US-09-989-732-414	Sequence 414, App
705	6	2.2	263	15	US-10-127-821A-484	Sequence 484, App	778	6	2.2	313	10	US-09-989-732-414	Sequence 414, App
706	6	2.2	263	15	US-10-127-822A-484	Sequence 484, App	779	6	2.2	313	10	US-09-991-073-414	Sequence 414, App
707	6	2.2	263	15	US-10-127-824A-484	Sequence 484, App	780	6	2.2	313	10	US-09-990-442-414	Sequence 414, App
708	6	2.2	263	15	US-10-127-826A-484	Sequence 484, App	781	6	2.2	313	10	US-09-991-163-414	Sequence 414, App
709	6	2.2	263	15	US-10-127-827A-484	Sequence 484, App	782	6	2.2	313	10	US-09-993-604-414	Sequence 414, App
710	6	2.2	263	15	US-10-127-828A-484	Sequence 484, App	783	6	2.2	313	10	US-09-990-444-414	Sequence 414, App
711	6	2.2	263	15	US-10-127-830A-484	Sequence 484, App	784	6	2.2	313	10	US-09-989-721-414	Sequence 414, App
712	6	2.2	263	15	US-10-127-832A-484	Sequence 484, App	785	6	2.2	313	10	US-09-992-599-414	Sequence 414, App
713	6	2.2	263	15	US-10-127-833A-484	Sequence 484, App	786	6	2.2	313	10	US-09-989-299A-414	Sequence 414, App
714	6	2.2	263	15	US-10-127-834A-484	Sequence 484, App	787	6	2.2	313	10	US-09-989-733-414	Sequence 414, App
715	6	2.2	263	15	US-10-127-836A-484	Sequence 484, App	788	6	2.2	313	10	US-09-990-444-414	Sequence 414, App
716	6	2.2	263	15	US-10-127-841A-484	Sequence 484, App	789	6	2.2	313	10	US-09-990-444-414	Sequence 414, App
717	6	2.2	263	15	US-10-127-844A-484	Sequence 484, App	790	6	2.2	313	10	US-09-991-181-414	Sequence 414, App
718	6	2.2	263	15	US-10-128-687A-484	Sequence 484, App	791	6	2.2	313	10	US-09-989-730-414	Sequence 414, App
719	6	2.2	263	15	US-10-128-688A-484	Sequence 484, App	792	6	2.2	313	10	US-09-990-436-414	Sequence 414, App
720	6	2.2	263	15	US-10-128-689A-484	Sequence 484, App	793	6	2.2	313	11	US-09-993-687-414	Sequence 414, App
721	6	2.2	263	15	US-10-128-694A-484	Sequence 484, App	794	6	2.2	313	11	US-09-989-734-414	Sequence 414, App
722	6	2.2	263	15	US-10-131-825A-484	Sequence 484, App	795	6	2.2	313	11	US-09-997-653-414	Sequence 414, App
723	6	2.2	263	15	US-10-130-417-484	Sequence 484, App	796	6	2.2	313	11	US-09-993-657-414	Sequence 414, App
724	6	2.2	263	15	US-10-131-815A-484	Sequence 484, App	797	6	2.2	313	11	US-09-997-428-414	Sequence 414, App
725	6	2.2	263	15	US-10-131-817A-484	Sequence 484, App	798	6	2.2	313	11	US-09-997-666-414	Sequence 414, App
726	6	2.2	263	15	US-10-131-821A-484	Sequence 484, App	799	6	2.2	313	11	US-09-990-433-414	Sequence 414, App
727	6	2.2	263	15	US-10-131-822A-484	Sequence 484, App	800	6	2.2	313	11	US-09-990-563-414	Sequence 414, App
728	6	2.2	263	15	US-10-131-828A-484	Sequence 484, App	801	6	2.2	313	11	US-09-990-711-414	Sequence 414, App
729	6	2.2	263	15	US-10-131-835A-484	Sequence 484, App	802	6	2.2	313	11	US-09-989-726-414	Sequence 414, App
730	6	2.2	263	15	US-10-137-864A-484	Sequence 484, App	803	6	2.2	313	11	US-09-998-156-414	Sequence 414, App
731	6	2.2	263	15	US-10-137-869A-484	Sequence 484, App	804	6	2.2	313	11	US-09-990-437-414	Sequence 414, App
732	6	2.2	263	15	US-10-147-523-484	Sequence 484, App	805	6	2.2	313	11	US-09-991-157-414	Sequence 414, App
733	6	2.2	263	15	US-10-158-785-484	Sequence 484, App	806	6	2.2	313	11	US-09-997-515-414	Sequence 414, App
734	6	2.2	263	15	US-10-121-051-484	Sequence 484, App	807	6	2.2	313	11	US-09-997-573-414	Sequence 414, App
735	6	2.2	263	15	US-10-121-042-484	Sequence 484, App	808	6	2.2	313	11	US-09-991-177-414	Sequence 414, App
736	6	2.2	263	15	US-10-123-912-484	Sequence 484, App	809	6	2.2	313	11	US-09-990-726-414	Sequence 414, App
737	6	2.2	263	15	US-10-223-085-182	Sequence 182, App	810	6	2.2	313	11	US-09-997-553-414	Sequence 414, App
738	6	2.2	263	15	US-10-192-007-484	Sequence 484, App	811	6	2.2	313	11	US-09-997-601-414	Sequence 414, App
739	6	2.2	263	15	US-10-194-359-484	Sequence 484, App	812	6	2.2	313	11	US-09-990-443-414	Sequence 414, App
740	6	2.2	263	15	US-10-226-739-41	Sequence 41, App1	813	6	2.2	313	11	US-09-991-854-414	Sequence 414, App
741	6	2.2	263	15	US-10-223-084-182	Sequence 182, App	814	6	2.2	313	11	US-09-997-628-414	Sequence 414, App
742	6	2.2	263	15	US-10-223-088-182	Sequence 182, App	815	6	2.2	313	11	US-09-997-652-414	Sequence 414, App
743	6	2.2	263	15	US-10-223-090-182	Sequence 182, App	816	6	2.2	313	11	US-09-989-728A-414	Sequence 414, App
744	6	2.2	263	15	US-10-223-087-182	Sequence 182, App	817	6	2.2	313	11	US-09-997-349-414	Sequence 414, App
745	6	2.2	263	15	US-10-127-847A-484	Sequence 484, App	818	6	2.2	313	11	US-09-997-440-414	Sequence 414, App

819	6	2.2	313	11	US-09-990-440-414	Sequence 414, App	892	6	2.2	313	12	US-10-063-514-88	Sequence 88, App
820	6	2.2	313	11	US-09-993-469-414	Sequence 414, App	893	6	2.2	313	12	US-10-063-512-88	Sequence 88, App
821	6	2.2	313	11	US-09-997-542-414	Sequence 414, App	894	6	2.2	313	12	US-10-063-521-88	Sequence 88, App
822	6	2.2	313	11	US-09-993-748-414	Sequence 414, App	895	6	2.2	313	12	US-10-063-527-88	Sequence 88, App
823	6	2.2	313	11	US-09-990-439-414	Sequence 414, App	896	6	2.2	313	12	US-10-063-528-88	Sequence 88, App
824	6	2.2	313	11	US-09-990-427-414	Sequence 414, App	897	6	2.2	313	12	US-10-063-522-88	Sequence 88, App
825	6	2.2	313	11	US-09-989-328-414	Sequence 414, App	898	6	2.2	313	12	US-10-063-530-88	Sequence 88, App
826	6	2.2	313	11	US-09-993-563-414	Sequence 414, App	899	6	2.2	313	12	US-10-063-546-88	Sequence 88, App
827	6	2.2	313	11	US-09-771-503-9	Sequence 9, Appl	900	6	2.2	313	12	US-10-063-556-88	Sequence 88, App
828	6	2.2	313	11	US-09-941-992-414	Sequence 414, App	901	6	2.2	313	12	US-10-063-562-88	Sequence 88, App
829	6	2.2	313	11	US-09-992-521-414	Sequence 414, App	902	6	2.2	313	12	US-10-063-566-88	Sequence 88, App
830	6	2.2	313	11	US-09-997-333-414	Sequence 414, App	903	6	2.2	313	12	US-10-063-568-88	Sequence 88, App
831	6	2.2	313	11	US-09-997-384-414	Sequence 414, App	904	6	2.2	313	12	US-10-063-568-88	Sequence 88, App
832	6	2.2	313	11	US-09-998-041-414	Sequence 414, App	905	6	2.2	313	12	US-10-063-570-88	Sequence 88, App
833	6	2.2	313	11	US-09-998-041-414	Sequence 414, App	906	6	2.2	313	12	US-10-063-577-88	Sequence 88, App
834	6	2.2	313	11	US-09-997-585-414	Sequence 414, App	907	6	2.2	313	12	US-10-063-579-88	Sequence 88, App
835	6	2.2	313	11	US-09-997-614-414	Sequence 414, App	908	6	2.2	313	12	US-10-063-581-88	Sequence 88, App
836	6	2.2	313	12	US-09-989-862-414	Sequence 414, App	909	6	2.2	313	12	US-10-063-582-88	Sequence 88, App
837	6	2.2	313	12	US-09-989-725-414	Sequence 414, App	910	6	2.2	313	12	US-10-063-584-88	Sequence 88, App
838	6	2.2	313	12	US-09-997-529-414	Sequence 414, App	911	6	2.2	313	12	US-10-063-588-88	Sequence 88, App
839	6	2.2	313	12	US-10-063-735-88	Sequence 88, Appl	912	6	2.2	313	12	US-10-063-587-88	Sequence 88, App
840	6	2.2	313	12	US-10-199-672-294	Sequence 294, App	913	6	2.2	313	12	US-10-063-589-88	Sequence 88, App
841	6	2.2	313	12	US-10-187-749-294	Sequence 294, App	914	6	2.2	313	12	US-10-063-592-88	Sequence 88, App
842	6	2.2	313	12	US-10-194-457-294	Sequence 294, App	915	6	2.2	313	12	US-10-063-592-88	Sequence 88, App
843	6	2.2	313	12	US-10-184-642-294	Sequence 294, App	916	6	2.2	313	12	US-10-063-592-88	Sequence 88, App
844	6	2.2	313	12	US-10-196-747-294	Sequence 294, App	917	6	2.2	313	12	US-10-063-593-88	Sequence 88, App
845	6	2.2	313	12	US-10-173-689-294	Sequence 294, App	918	6	2.2	313	12	US-10-063-596-88	Sequence 88, App
846	6	2.2	313	12	US-10-173-690-294	Sequence 294, App	919	6	2.2	313	12	US-10-063-600-88	Sequence 88, App
847	6	2.2	313	12	US-10-173-691-294	Sequence 294, App	920	6	2.2	313	12	US-10-063-602-88	Sequence 88, App
848	6	2.2	313	12	US-10-173-692-294	Sequence 294, App	921	6	2.2	313	12	US-10-063-604-88	Sequence 88, App
849	6	2.2	313	12	US-10-173-694-294	Sequence 294, App	922	6	2.2	313	12	US-10-063-606-88	Sequence 88, App
850	6	2.2	313	12	US-10-173-698-294	Sequence 294, App	923	6	2.2	313	12	US-10-063-607-88	Sequence 88, App
851	6	2.2	313	12	US-10-173-699-294	Sequence 294, App	924	6	2.2	313	12	US-10-063-610-88	Sequence 88, App
852	6	2.2	313	12	US-10-173-707-294	Sequence 294, App	925	6	2.2	313	12	US-10-063-609-88	Sequence 88, App
853	6	2.2	313	12	US-10-174-569-294	Sequence 294, App	926	6	2.2	313	12	US-10-063-611-88	Sequence 88, App
854	6	2.2	313	12	US-10-174-583-294	Sequence 294, App	927	6	2.2	313	12	US-10-063-612-88	Sequence 88, App
855	6	2.2	313	12	US-10-174-587-294	Sequence 294, App	928	6	2.2	313	12	US-10-063-614-88	Sequence 88, App
856	6	2.2	313	12	US-10-174-589-294	Sequence 294, App	929	6	2.2	313	12	US-10-063-615-88	Sequence 88, App
857	6	2.2	313	12	US-10-174-591-294	Sequence 294, App	930	6	2.2	313	12	US-10-063-618-88	Sequence 88, App
858	6	2.2	313	12	US-10-175-736-294	Sequence 294, App	931	6	2.2	313	12	US-10-063-623-88	Sequence 88, App
859	6	2.2	313	12	US-10-175-742-294	Sequence 294, App	932	6	2.2	313	12	US-10-063-640-88	Sequence 88, App
860	6	2.2	313	12	US-10-175-744-294	Sequence 294, App	933	6	2.2	313	12	US-10-063-644-88	Sequence 88, App
861	6	2.2	313	12	US-10-175-745-294	Sequence 294, App	934	6	2.2	313	12	US-10-063-644-88	Sequence 88, App
862	6	2.2	313	12	US-10-175-748-294	Sequence 294, App	935	6	2.2	313	12	US-10-063-644-88	Sequence 88, App
863	6	2.2	313	12	US-10-175-751-294	Sequence 294, App	936	6	2.2	313	12	US-10-063-644-88	Sequence 88, App
864	6	2.2	313	12	US-10-175-754-294	Sequence 294, App	937	6	2.2	313	12	US-10-063-646-88	Sequence 88, App
865	6	2.2	313	12	US-10-176-480-294	Sequence 294, App	938	6	2.2	313	12	US-10-063-650-88	Sequence 88, App
866	6	2.2	313	12	US-10-176-489-294	Sequence 294, App	939	6	2.2	313	12	US-10-063-651-88	Sequence 88, App
867	6	2.2	313	12	US-10-176-754-294	Sequence 294, App	940	6	2.2	313	12	US-10-063-651-88	Sequence 88, App
868	6	2.2	313	12	US-10-176-755-294	Sequence 294, App	941	6	2.2	313	12	US-10-063-654-88	Sequence 88, App
869	6	2.2	313	12	US-10-176-759-294	Sequence 294, App	942	6	2.2	313	12	US-10-063-654-88	Sequence 88, App
870	6	2.2	313	12	US-10-176-920-294	Sequence 294, App	943	6	2.2	313	12	US-10-063-659-88	Sequence 88, App
871	6	2.2	313	12	US-10-176-922-294	Sequence 294, App	944	6	2.2	313	12	US-10-063-660-88	Sequence 88, App
872	6	2.2	313	12	US-10-176-924-294	Sequence 294, App	945	6	2.2	313	12	US-10-063-661-88	Sequence 88, App
873	6	2.2	313	12	US-10-176-984-294	Sequence 294, App	946	6	2.2	313	12	US-10-063-662-88	Sequence 88, App
874	6	2.2	313	12	US-10-179-508-294	Sequence 294, App	947	6	2.2	313	12	US-10-063-665-88	Sequence 88, App
875	6	2.2	313	12	US-10-179-512-294	Sequence 294, App	948	6	2.2	313	12	US-10-063-666-88	Sequence 88, App
876	6	2.2	313	12	US-10-179-515-294	Sequence 294, App	949	6	2.2	313	12	US-10-063-668-88	Sequence 88, App
877	6	2.2	313	12	US-10-063-526-88	Sequence 88, Appl	950	6	2.2	313	12	US-10-063-670-88	Sequence 88, App
878	6	2.2	313	12	US-10-173-702-294	Sequence 294, App	951	6	2.2	313	12	US-10-063-671-88	Sequence 88, App
879	6	2.2	313	12	US-10-173-703-294	Sequence 294, App	952	6	2.2	313	12	US-10-063-672-88	Sequence 88, App
880	6	2.2	313	12	US-10-173-704-294	Sequence 294, App	953	6	2.2	313	12	US-10-063-674-88	Sequence 88, App
881	6	2.2	313	12	US-10-174-574-294	Sequence 294, App	954	6	2.2	313	12	US-10-063-677-88	Sequence 88, App
882	6	2.2	313	12	US-10-176-486-294	Sequence 294, App	955	6	2.2	313	12	US-10-063-678-88	Sequence 88, App
883	6	2.2	313	12	US-10-176-490-294	Sequence 294, App	956	6	2.2	313	12	US-10-063-676-88	Sequence 88, App
884	6	2.2	313	12	US-10-176-752-294	Sequence 294, App	957	6	2.2	313	12	US-10-063-682-88	Sequence 88, App
885	6	2.2	313	12	US-10-176-981-294	Sequence 294, App	958	6	2.2	313	12	US-10-063-686-88	Sequence 88, App
886	6	2.2	313	12	US-10-176-983-294	Sequence 294, App	959	6	2.2	313	12	US-10-063-688-88	Sequence 88, App
887	6	2.2	313	12	US-10-176-988-294	Sequence 294, App	960	6	2.2	313	12	US-10-063-692-88	Sequence 88, App
888	6	2.2	313	12	US-10-179-517-294	Sequence 294, App	961	6	2.2	313	12	US-10-063-693-88	Sequence 88, App
889	6	2.2	313	12	US-10-179-521-294	Sequence 294, App	962	6	2.2	313	12	US-10-063-694-88	Sequence 88, App
890	6	2.2	313	12	US-10-063-586-88	Sequence 88, Appl	963	6	2.2	313	12	US-10-063-698-88	Sequence 88, App
891	6	2.2	313	12	US-10-063-510-88	Sequence 88, Appl	964	6	2.2	313	12	US-10-063-699-88	Sequence 88, App

```
965 6 2.2 313 12 US-10-063-702-88 Sequence 88, Appl
966 6 2.2 313 12 US-10-063-705-88 Sequence 88, Appl
967 6 2.2 313 12 US-10-063-707-88 Sequence 88, Appl
968 6 2.2 313 12 US-10-063-709-88 Sequence 88, Appl
969 6 2.2 313 12 US-10-063-710-88 Sequence 88, Appl
970 6 2.2 313 12 US-10-063-711-88 Sequence 88, Appl
971 6 2.2 313 12 US-10-063-712-88 Sequence 88, Appl
972 6 2.2 313 12 US-10-063-713-88 Sequence 88, Appl
973 6 2.2 313 12 US-10-063-714-88 Sequence 88, Appl
974 6 2.2 313 12 US-10-063-715-88 Sequence 88, Appl
975 6 2.2 313 12 US-10-063-716-88 Sequence 88, Appl
976 6 2.2 313 12 US-10-063-717-88 Sequence 88, Appl
977 6 2.2 313 12 US-10-063-720-88 Sequence 88, Appl
978 6 2.2 313 12 US-10-063-721-88 Sequence 88, Appl
979 6 2.2 313 12 US-10-063-722-88 Sequence 88, Appl
980 6 2.2 313 12 US-10-063-723-88 Sequence 88, Appl
981 6 2.2 313 12 US-10-063-724-88 Sequence 88, Appl
982 6 2.2 313 12 US-10-063-726-88 Sequence 88, Appl
983 6 2.2 313 12 US-10-063-727-88 Sequence 88, Appl
984 6 2.2 313 12 US-10-063-728-88 Sequence 88, Appl
985 6 2.2 313 12 US-10-063-730-88 Sequence 88, Appl
986 6 2.2 313 12 US-10-063-731-88 Sequence 88, Appl
987 6 2.2 313 12 US-10-063-732-88 Sequence 88, Appl
988 6 2.2 313 12 US-10-063-734-88 Sequence 88, Appl
989 6 2.2 313 12 US-10-063-736-88 Sequence 88, Appl
990 6 2.2 313 12 US-10-063-742-88 Sequence 88, Appl
991 6 2.2 313 12 US-10-063-743-88 Sequence 88, Appl
992 6 2.2 313 12 US-10-063-744-88 Sequence 88, Appl
993 6 2.2 313 12 US-10-063-745-294 Sequence 294, App
994 6 2.2 313 12 US-10-063-520-88 Sequence 88, Appl
995 6 2.2 313 12 US-10-063-548-88 Sequence 88, Appl
996 6 2.2 313 12 US-10-063-578-88 Sequence 88, Appl
997 6 2.2 313 12 US-10-063-647-88 Sequence 88, Appl
998 6 2.2 313 12 US-10-063-648-88 Sequence 88, Appl
999 6 2.2 313 12 US-10-063-677-88 Sequence 88, Appl
1000 6 2.2 313 12 US-10-063-684-88 Sequence 88, Appl
```

## ALIGNMENTS

```
RESULT 1
; Sequence 2, Application US/10033190
; Publication No. US20020133848A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis Plant Sciences, Inc.
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF AN ANTHOCYANIN MUTANT (ANT)
; FILE REFERENCE: EP01-002C
; CURRENT APPLICATION NUMBER: US/10/033,190
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,685
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
; US-10-033-190-2
```

```
Query Match 100.0%; Score 274; DB 14; Length 274;
Best Local Similarity 100.0%; Pred. No. 3.8e-265;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MNSTMSLGVKSGWTDDEDFLLKRCIDKYGSGKWLVPFAGLNRCSRRLRWNTYL 60
DB 1 MNSTMSLGVKSGWTDDEDFLLKRCIDKYGSGKWLVPFAGLNRCSRRLRWNTYL 60
QY 61 RPHIKRGDFODEVDLLRLHLKLLGNRWSLIAGRLPGRTANDVKNYNTNLLRLKNTTKI 120
DB 61 RPHIKRGDFODEVDLLRLHLKLLGNRWSLIAGRLPGRTANDVKNYNTNLLRLKNTTKI 120
```

```
QY 121 VREKINNKGEISTKIEIKPQRRKYFSTMKNTNNVILDEEHCKEIISEKOTPPA 180
DB 121 VREKINNKGEISTKIEIKPQRRKYFSTMKNTNNVILDEEHCKEIISEKOTPPA 180
QY 181 SMDNVDPWMINLLENCNDIEDEEVVINYKTLTSLHBEISPPINIGSGNSMOGOIS 240
DB 181 SMDNVDPWMINLLENCNDIEDEEVVINYKTLTSLHBEISPPINIGSGNSMOGOIS 240
QY 241 HENWGEFSINLPMPQGVQNDPFSAEIDILMNLID 274
DB 241 HENWGEFSINLPMPQGVQNDPFSAEIDILMNLID 274
```

```
RESULT 2
; Sequence 480, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Rieckmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Plineda, Omaira
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 480
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-09-934-455-480
```

```
Query Match 11.7%; Score 32; DB 11; Length 209;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 78 LRLHLKLLGNRWSLIAGRLPGRTANDVKNYNT 109
DB 75 LRLHLKLLGNRWSLIAGRLPGRTANDVKNYNT 106
```

```
RESULT 3
; Sequence 482, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Rieckmann, Jose Luis
```

```

; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omatira
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 482
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-09-934-455-482

```

```

Query Match 11.7%; Score 32; DB 11; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 78 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109
DB 75 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 106

```

```

RESULT 4
US-09-934-455-14
; Sequence 14, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omatira
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-09-934-455-14

```

```

Query Match 11.7%; Score 32; DB 11; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 78 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109
DB 75 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 106

```

```

RESULT 5
US-10-295-403-44
; Sequence 44, Application US/10295403
; Publication No. US20030101481A1
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc
; APPLICANT: Brown, Pierre
; APPLICANT: Pineda, Omatira
; APPLICANT: Reuber, Lynne
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Zhang, James
; APPLICANT: Benito, Maria-Ines
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Fromm, Mike
; TITLE OF INVENTION: PLANT GENE SEQUENCES I
; FILE REFERENCE: MBI-0003
; CURRENT APPLICATION NUMBER: US/10/295,403
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 60/101,349
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: 60/103,312
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/108,734
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/113,409
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G663
; US-10-295-403-44

```

```

Query Match 11.7%; Score 32; DB 15; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 78 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109
DB 75 LRLHKLGNRWSLIAGRLPGRTANDVKNYNT 106

```

```

RESULT 6
US-09-934-455-438
; Sequence 438, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omatira
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22

```

PRIOR APPLICATION NUMBER: MBI-0022  
PRIOR FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: MBI-0023  
PRIOR FILING DATE: 2001-04-17  
NUMBER OF SEQ ID NOS: 516  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 438  
LENGTH: 212  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-934-455-438

Query Match 9.9%; Score 27; DB 11; Length 212;  
Best Local Similarity 100.0%; Pred. No. 1.2e-18;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 LIGNRWSLIAGRLPGRTANDVKYMYNT 109  
DB 80 LIGNRWSLIAGRLPGRTANDVKYMYNT 106

RESULT 7  
US-09-443-704-6  
Sequence 6, Application US/09443704  
Patent No. US20020066120A1  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Liu, Zhan-Bin  
APPLICANT: Odell, Joan  
APPLICANT: Rafalecki, Antoni  
APPLICANT: Shi, June  
APPLICANT: Meng, Zude  
TITLE OF INVENTION: Plant MYB-Related Transcription Factors  
FILE REFERENCE: BBI280 US NA  
CURRENT APPLICATION NUMBER: US/09/443,704  
CURRENT FILING DATE: 1999-11-19  
EARLIER APPLICATION NUMBER: 60/109,294  
EARLIER FILING DATE: No. US20020066120A1ember 20, 1998  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 6  
LENGTH: 120  
TYPE: PRT  
ORGANISM: Glycine max  
US-09-443-704-6

Query Match 6.2%; Score 17; DB 9; Length 120;  
Best Local Similarity 100.0%; Pred. No. 7.2e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 LIGNRWSLIAGRLPGRT 99  
DB 84 LIGNRWSLIAGRLPGRT 100

RESULT 8  
US-10-008-118A-6  
Sequence 6, Application US/10008118A  
Publication No. US20020187539A1  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Meng, Zude  
TITLE OF INVENTION: Plant MYB-Related Transcription Factors  
FILE REFERENCE: BBI280 USDIY  
CURRENT APPLICATION NUMBER: US/10/008,118A  
CURRENT FILING DATE: 2001-12-05  
PRIOR APPLICATION NUMBER: 60/109,294  
PRIOR FILING DATE: 1998-11-20  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 6  
LENGTH: 120  
TYPE: PRT

ORGANISM: Glycine max  
US-10-008-118A-6

Query Match 6.2%; Score 17; DB 14; Length 120;  
Best Local Similarity 100.0%; Pred. No. 7.2e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 LIGNRWSLIAGRLPGRT 99  
DB 84 LIGNRWSLIAGRLPGRT 100

RESULT 9  
US-09-934-455-390  
Sequence 390, Application US/09934455  
Publication No. US20030121070A1  
GENERAL INFORMATION:  
APPLICANT: Adam, Luc  
APPLICANT: Creelman, Robert  
APPLICANT: Dubell, Arnold  
APPLICANT: Heard, Jacqueline  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Keddie, James  
APPLICANT: Pilgrim, Marsha  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Reuber, Lynne  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Yu, Guo-Liang  
APPLICANT: Pineda, Omaira  
TITLE OF INVENTION: Genes for Modifying Plant Traits IV  
FILE REFERENCE: MBI-0025  
CURRENT APPLICATION NUMBER: US/09/934,455  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/227439  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: MBI-0022  
PRIOR FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: MBI-0023  
PRIOR FILING DATE: 2001-04-17  
NUMBER OF SEQ ID NOS: 516  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 390  
LENGTH: 203  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-934-455-390

Query Match 5.1%; Score 14; DB 11; Length 203;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 RLHKLIGNRWSLIA 92  
DB 84 RLHKLIGNRWSLIA 97

RESULT 10  
US-10-407-920-29  
Sequence 29, Application US/10407920  
Publication No. US20040006797A1  
GENERAL INFORMATION:  
APPLICANT: Shi, Lifang  
TITLE OF INVENTION: MYB Transcription Factors and Uses for Crop Improvement  
FILE REFERENCE: 38-21(52703)A  
CURRENT APPLICATION NUMBER: US/10/407,920  
CURRENT FILING DATE: 2003-06-06  
PRIOR APPLICATION NUMBER: US/60/370,759  
PRIOR FILING DATE: 2002-04-10  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 29  
LENGTH: 203  
TYPE: PRT

ORGANISM: Arabidopsis thaliana  
US-10-407-920-29

Query Match  
Best Local Similarity 100.0%; Score 14; DB 12; Length 203;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 RLHKLGNRWSLIA 92  
DB 84 RLHKLGNRWSLIA 97

RESULT 11  
US-09-934-455-70

Sequence 70, Application US/09934455  
Publication No. US20030121070A1

GENERAL INFORMATION:

APPLICANT: Adam, Luc  
APPLICANT: Creelman, Robert  
APPLICANT: Dubell, Arnold  
APPLICANT: Heard, Jacqueline  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Keddie, James  
APPLICANT: Pilgrim, Marsha  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Reuber, Lynne  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: Genes for Modifying Plant Traits IV  
FILE REFERENCE: MBI-0025  
CURRENT APPLICATION NUMBER: US/09/934,455  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/227439  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: MBI-0022  
PRIOR FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: MBI-0023  
PRIOR FILING DATE: 2001-04-17  
NUMBER OF SEQ ID NOS: 516  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 70  
LENGTH: 219  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-934-455-70

Query Match  
Best Local Similarity 100.0%; Score 14; DB 11; Length 219;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 RLHKLGNRWSLIA 92  
DB 80 RLHKLGNRWSLIA 93

RESULT 12

US-10-021-811-22

Sequence 22, Application US/10021811  
Publication No. US20030024007A1

GENERAL INFORMATION:

APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Fang, Yiwen  
APPLICANT: Odell, Joan  
APPLICANT: Weng, Zude  
TITLE OF INVENTION: Plant Myb Transcription Factor Homologs  
FILE REFERENCE: BB1294 US NA  
CURRENT APPLICATION NUMBER: US/10/021,811  
CURRENT FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: 60/110,609  
PRIOR FILING DATE: 1998-December-02  
NUMBER OF SEQ ID NOS: 63  
SOFTWARE: Microsoft Office 97

SEQ ID NO 22  
LENGTH: 115  
TYPE: PRT  
ORGANISM: Oryza sativa

FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (53)  
NAME/KEY: UNSURE  
LOCATION: (100)  
NAME/KEY: UNSURE  
LOCATION: (112)  
US-10-021-811-22

Query Match  
Best Local Similarity 100.0%; Score 13; DB 15; Length 115;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 KSCRLRWLNYLRP 62  
DB 59 KSCRLRWLNYLRP 71

RESULT 13

US-10-021-811-6

Sequence 6, Application US/10021811  
Publication No. US20030024007A1

GENERAL INFORMATION:

APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Fang, Yiwen  
APPLICANT: Odell, Joan  
APPLICANT: Weng, Zude  
TITLE OF INVENTION: Plant Myb Transcription Factor Homologs  
FILE REFERENCE: BB1294 US NA  
CURRENT APPLICATION NUMBER: US/10/021,811  
CURRENT FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: 60/110,609  
PRIOR FILING DATE: 1998-December-02  
NUMBER OF SEQ ID NOS: 63  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 6  
LENGTH: 120  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (101)  
NAME/KEY: UNSURE  
LOCATION: (113) ..(114)  
US-10-021-811-6

Query Match  
Best Local Similarity 100.0%; Score 13; DB 15; Length 120;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 KSCRLRWLNYLRP 62  
DB 61 KSCRLRWLNYLRP 73

RESULT 14

US-10-021-811-60

Sequence 60, Application US/10021811  
Publication No. US20030024007A1

GENERAL INFORMATION:

APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Fang, Yiwen  
APPLICANT: Odell, Joan  
APPLICANT: Weng, Zude  
TITLE OF INVENTION: Plant Myb Transcription Factor Homologs  
FILE REFERENCE: BB1294 US NA  
CURRENT APPLICATION NUMBER: US/10/021,811  
CURRENT FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: 60/110,609

```

; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 60
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (27)
; NAME/KEY: UNSURE
; LOCATION: (109)
; NAME/KEY: UNSURE
; LOCATION: (111)
; NAME/KEY: UNSURE
; LOCATION: (122)
; NAME/KEY: UNSURE
; LOCATION: (129)
; US-10-021-811-60

```

Query Match 4.7%; Score 13; DB 15; Length 131;

Best Local Similarity 100.0%; Pred. No. 7.6e-05; Mismatches 0; Indels 0; Gaps 0;

```

Qy 50 KSCRLRWLNYLRP 62
Db 68 KSCRLRWLNYLRP 80

```

```

RESULT 15
US-10-021-811-18
; Sequence 18, Application US/10021811
; Publication No US20030024007A1
; GENERAL INFORMATION:
; APPLICANT: Canoon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Odell, Joan
; APPLICANT: Meng, Zude
; TITLE OF INVENTION: Plant W/b Transcription Factor Homologs
; FILE REFERENCE: BBI294 US NA
; CURRENT APPLICATION NUMBER: US/10/021,811
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-021-811-18

```

Query Match 4.7%; Score 13; DB 15; Length 145;

Best Local Similarity 100.0%; Pred. No. 8.3e-05; Mismatches 0; Indels 0; Gaps 0;

```

Qy 50 KSCRLRWLNYLRP 62
Db 58 KSCRLRWLNYLRP 70

```

Search completed: January 29, 2004, 20:26:32  
Job time : 47 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

## OM protein - protein search, using sw model

Run on: January 29, 2004, 20:17:42 : Search time 20 Seconds  
(without alignments)  
1317.511 Million cell updates/sec

Title: US-10-033-190-2

Perfect score: 274

Sequence: 1 NMSTWSMSLGVKSGSWTDE.....QQGVNDPFSARIDMLND 274

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

PIR 76:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	11.7	139	2	B96689
2	32	11.7	246	2	A96689
3	32	11.7	249	2	T51687
4	27	9.9	212	2	T51680
5	27	9.9	248	2	B96680
6	25	9.1	193	2	T09744
7	21	7.7	198	2	T09743
8	21	7.7	256	2	T01189
9	21	7.7	264	2	T03974
10	21	7.7	266	2	T03972
11	21	7.7	267	2	T03188
12	21	7.7	271	2	T03715
13	19	6.9	130	2	T51684
14	17	6.2	195	2	F96734
15	17	6.2	213	1	S71287
16	17	6.2	226	2	T09773
17	17	6.2	273	1	TVZMMB
18	17	6.2	275	2	T02987
19	17	6.2	321	2	T02987
20	14	5.1	128	2	T51672
21	14	5.1	203	1	T51420
22	14	5.1	228	1	TVMUG1
23	13	4.7	198	1	J00959
24	13	4.7	205	1	J00958
25	13	4.7	216	2	D96520
26	13	4.7	217	2	T06455
27	13	4.7	268	2	T03729
28	13	4.7	271	2	B96773
29	13	4.7	273	1	J02390

30	13	4.7	296	2	T06025	transcription fact
31	13	4.7	298	2	H84785	probable MYB fam11
32	13	4.7	310	2	T46035	AtMYB84 - Arabidop
33	13	4.7	333	2	T51650	probable transcrip
34	13	4.7	333	2	T09736	myb-related protei
35	13	4.7	333	2	T09735	ABA-induced myb-re
36	13	4.7	335	2	T09731	dehydration-induce
37	13	4.7	335	2	T05891	myb-related protei
38	13	4.7	376	2	T51673	myb-related transc
39	12	4.4	332	1	J00960	myb-related protei
40	12	4.4	236	2	D85096	probable DNA-bindi
41	12	4.4	246	2	T51631	probable transcrip
42	12	4.4	249	1	S68688	myb-related protei
43	12	4.4	257	2	T00780	myb-related protei
44	12	4.4	267	1	S61506	myb-related protei
45	12	4.4	273	1	S69189	myb-related protei
46	12	4.4	276	2	T02985	myb-related protei
47	12	4.4	282	2	T05690	myb-related transc
48	12	4.4	282	2	T51632	myb-related transc
49	12	4.4	294	2	T09879	myb-related protei
50	11	4.0	267	2	H86320	probable MYB47 tra
51	11	4.0	278	2	C86383	probable Myb-like
52	11	4.0	286	2	H96706	probable transcrip
53	10	3.6	45	2	T52283	R2R3-MYB transcrip
54	10	3.6	45	2	T52276	R2R3-MYB transcrip
55	10	3.6	239	2	T51686	probable transcrip
56	10	3.6	243	2	T51652	myb-related transc
57	10	3.6	293	2	T09758	myb-related protei
58	10	3.6	371	2	T00438	probable MYB fam11
59	10	3.6	371	2	T51636	myb-related transc
60	9	3.3	45	2	T52127	Myb-like transcrip
61	9	3.3	376	2	T03988	myb-like transcrip
62	9	3.3	399	1	A39697	maize myb-related
63	9	3.3	453	2	T09745	myb-related protei
64	8	2.9	214	2	T51657	myb-related transc
65	8	2.9	220	2	T51648	probable transcrip
66	8	2.9	434	2	G83737	C4-dicarboxylate t
67	8	2.9	1607	2	T52022	MAP kinase
68	7	2.6	45	2	T52271	R2R3-MYB transcrip
69	7	2.6	45	2	T52278	R2R3-MYB transcrip
70	7	2.6	67	2	T03833	myb7 protein - ric
71	7	2.6	122	1	S58294	myb-related protei
72	7	2.6	125	2	B75603	hypothetical prote
73	7	2.6	147	2	E97039	hypothetical prote
74	7	2.6	151	1	B39697	myb-related protei
75	7	2.6	153	2	T04563	myb-related protei
76	7	2.6	163	2	T03753	hypothetical prote
77	7	2.6	208	2	A64089	nitrate/nitrite re
78	7	2.6	209	2	AH0369	nitrate/nitrite re
79	7	2.6	215	2	A40584	nitrate/nitrite re
80	7	2.6	215	2	AD0787	nitrate/nitrite re
81	7	2.6	215	2	B91014	nitrate/nitrite re
82	7	2.6	215	2	D85858	nitrate/nitrite re
83	7	2.6	218	2	C81104	transcription regu
84	7	2.6	234	2	T51666	myb-related transc
85	7	2.6	238	2	T46166	MYB27 protein - Ar
86	7	2.6	239	2	T02984	myb-related protei
87	7	2.6	243	2	G96760	probable myb-like
88	7	2.6	243	2	T51662	myb-related transc
89	7	2.6	246	1	S71283	myb-related protei
90	7	2.6	246	1	D86197	hypothetical prote
91	7	2.6	248	2	T18315	hypothetical prote
92	7	2.6	248	2	G86314	F2H15.17 protein -
93	7	2.6	249	2	B84717	probable MYB fam11
94	7	2.6	249	2	T51660	myb-related transc
95	7	2.6	250	2	T05996	hypothetical prote
96	7	2.6	250	2	T48145	probable transcrip
97	7	2.6	251	1	S35729	myb-related protei
98	7	2.6	255	1	S04899	myb-related protei
99	7	2.6	256	2	T49254	Myb DNA binding pr
100	7	2.6	257	2	T03825	myb protein homolo
101	7	2.6	261	2	T06650	myb-related protei
102	7	2.6	261	2	T46138	myb-like protein -

103	7	2.6	263	2	T48607	probable transcrip	176	7	2.6	412	2	T52561	sepin homolog Spn
104	7	2.6	265	1	T07395	myb-related transc	177	7	2.6	418	2	A86229	hypothetical prote
105	7	2.6	268	1	T00961	myb-related protei	178	7	2.6	414	2	B86171	hypothetical prote
106	7	2.6	269	1	S58282	probable MYB fami	179	7	2.6	421	1	S26605	myb-related protei
107	7	2.6	274	1	T00957	myb-related protei	180	7	2.6	421	1	S24244	myb-related protei
108	7	2.6	274	1	T05769	myb-related protei	181	7	2.6	422	1	W2WLB2	E2 protein - Bovin
109	7	2.6	274	2	D86300	hypothetical prote	182	7	2.6	427	2	A84667	probable MYB fami
110	7	2.6	274	2	T07393	myb-related transc	183	7	2.6	475	2	B64120	killing factor Kic
111	7	2.6	276	2	T03823	probable myb-relat	184	7	2.6	490	2	T02545	probable MYB fami
112	7	2.6	278	2	T03850	myb-related protei	185	7	2.6	502	1	B40620	pyruvate kinase (E
113	7	2.6	278	2	T51641	myb-related transc	186	7	2.6	502	2	D86791	pyruvate kinase (E
114	7	2.6	279	2	T03830	probable myb facto	187	7	2.6	514	2	T29652	inward rectifier p
115	7	2.6	280	1	S26604	myb-related protei	188	7	2.6	517	1	S26606	myb-related protei
116	7	2.6	280	2	T00737	myb-related protei	189	7	2.6	524	2	S33640	homeotic protein s
117	7	2.6	280	2	T51667	myb-related transc	190	7	2.6	527	2	T21830	hypothetical prote
118	7	2.6	280	2	T49901	probable transcrip	191	7	2.6	531	2	T01516	SRP1 protein homol
119	7	2.6	282	2	C96687	hypothetical prote	192	7	2.6	531	2	T52099	probable nuclear t
120	7	2.6	282	2	B85327	probable transcrip	193	7	2.6	531	2	B84492	probable retroelem
121	7	2.6	282	2	T24693	hypothetical prote	194	7	2.6	553	2	T06179	myb-related protei
122	7	2.6	282	2	D85295	myb-like protein (	195	7	2.6	553	2	T03762	hypothetical transc
123	7	2.6	283	2	T05791	myb-related protei	196	7	2.6	669	2	E71127	ATP-dependent heli
124	7	2.6	288	2	T05954	transcription fact	197	7	2.6	702	2	A69140	ATP-dependent heli
125	7	2.6	296	2	A96603	probable Myb-fam1	198	7	2.6	714	2	AD2236	p-aminobenzoic aci
126	7	2.6	297	2	T47857	myb protein-like -	199	7	2.6	745	2	A71448	probable MYB trans
127	7	2.6	299	2	T47917	probable transcrip	200	7	2.6	754	2	S62561	ATP dependent RNA
128	7	2.6	301	2	T45859	R2R3-MYB transcrip	201	7	2.6	779	2	S64680	exoribonuclease 10
129	7	2.6	302	1	S31818	myb-related protei	202	7	2.6	791	2	T41573	hypothetical prote
130	7	2.6	304	1	S71285	myb-related protei	203	7	2.6	844	2	T33608	hypothetical prote
131	7	2.6	305	1	S71284	myb-related protei	204	7	2.6	866	2	H86288	hypothetical prote
132	7	2.6	309	2	T00503	probable MYB fam1	205	7	2.6	938	2	G70472	hypothetical prote
133	7	2.6	311	2	T03827	myb protein homolo	206	7	2.6	956	2	B71468	probable insulinas
134	7	2.6	314	2	E96609	probable DNA-bind	207	7	2.6	1000	2	T39423	hypothetical prote
135	7	2.6	315	2	S66038	YacC protein - Bac	208	7	2.6	1156	2	B70356	chromosome assembl
136	7	2.6	316	1	U00956	myb-related protei	209	7	2.6	1200	2	T48194	hypothetical prote
137	7	2.6	320	1	C85440	myb-related protei	210	7	2.6	1204	2	A96676	hypothetical prote
138	7	2.6	321	1	S45338	myb-related protei	211	7	2.6	1245	2	D71613	GAF domain protein
139	7	2.6	323	2	T51621	myb-like protein (	212	7	2.6	1297	2	T52065	probable myb-relat
140	7	2.6	323	2	T51645	myb-related transc	213	7	2.6	1451	2	S65571	pattern formation
141	7	2.6	324	2	B85064	MYB-like protein (	214	7	2.6	1883	2	T13944	chromodomain helic
142	7	2.6	325	2	T51509	probable transcrip	215	7	2.6	2352	2	C83329	probable non-ribos
143	7	2.6	326	2	T49966	myb-related protei	216	7	2.2	30	2	S34765	4-hydroxybutyryl-C
144	7	2.6	327	2	T01038	myb-related protei	217	7	2.2	45	2	T07431	hypothetical prote
145	7	2.6	330	2	C96775	hypothetical prote	218	7	2.2	45	2	T52123	R2R3-MYB transcrip
146	7	2.6	330	2	C96717	hypothetical prote	219	7	2.2	45	2	T52272	R2R3-MYB transcrip
147	7	2.6	331	1	S69190	myb-related protei	220	7	2.2	45	2	T52277	R2R3-MYB transcrip
148	7	2.6	332	1	S58283	myb-related protei	221	7	2.2	45	2	T52282	R2R3-MYB transcrip
149	7	2.6	333	2	T45720	probable transcrip	222	7	2.2	57	2	S66314	protein kinase AKI
150	7	2.6	333	2	G96768	protein transcript	223	7	2.2	64	2	A99766	hypothetical prote
151	7	2.6	334	2	G01650	malate dehydrogena	224	7	2.2	64	2	S25439	beta-fructofuranos
152	7	2.6	334	2	T50816	probable transcrip	225	7	2.2	81	2	AF3471	hypothetical prote
153	7	2.6	336	2	F85021	probable transcrip	226	7	2.2	87	2	T18146	hypothetical prote
154	7	2.6	336	1	T51644	probable transcrip	227	7	2.2	87	2	T23257	hypothetical prote
155	7	2.6	340	1	S04898	myb-related protei	228	7	2.2	88	2	F69528	hypothetical prote
156	7	2.6	343	2	T48050	probable transcrip	229	7	2.2	96	2	A05306	beta-fructofuranos
157	7	2.6	343	2	T52590	myb-related transc	230	7	2.2	98	2	A64088	hypothetical prote
158	7	2.6	347	2	T07398	myb-related transc	231	7	2.2	101	2	T38892	very hypothetical
159	7	2.6	348	2	T02006	transcription fact	232	7	2.2	101	2	A13410	2-isopropylmalate
160	7	2.6	350	2	G84725	hypothetical prote	233	7	2.2	103	2	F75334	probable dioxigena
161	7	2.6	352	1	S58293	myb-related protei	234	7	2.2	112	2	F83564	probable HTT fami
162	7	2.6	352	2	T51659	myb-related protei	235	7	2.2	113	2	D72454	hypothetical prote
163	7	2.6	352	2	D86476	myb-related transc	236	7	2.2	117	2	AD0624	probable bacteriop
164	7	2.6	357	2	D86475	hypothetical prote	237	7	2.2	118	1	E64062	diacylglycerol kin
165	7	2.6	359	2	D86394	protein T24P13.16	238	7	2.2	118	2	T51664	myb-related transc
166	7	2.6	365	2	D86470	F21H2.9 protein -	239	7	2.2	119	2	A72665	hypothetical prote
167	7	2.6	367	2	T01017	probable MYB fam1	240	7	2.2	121	2	AE1264	nitrogen regulator
168	7	2.6	368	2	T03828	myb protein - rice	241	7	2.2	121	2	AG1626	nitrogen regulator
169	7	2.6	369	2	T01196	transcription fact	242	7	2.2	121	2	AC3021	conserved hypobet
170	7	2.6	370	2	C96664	hypothetical prote	243	7	2.2	121	2	F98263	hypothetical prote
171	7	2.6	386	2	B97035	uncharacterized pr	244	7	2.2	124	2	H72801	gpi9 protein - Myc
172	7	2.6	389	2	T04816	myb-related protei	245	7	2.2	125	2	T32914	hypothetical prote
173	7	2.6	393	1	S22520	myb-related protei	246	7	2.2	128	2	AG3404	hypothetical cytos
174	7	2.6	395	2	T02989	myb-related protei	247	7	2.2	131	2	AD0744	probable membrane
175	7	2.6	399	2	T47712	MYB transcrip	248	7	2.2	134	2	D70162	ribosomal protein

249	6	2.2	136	2	A87635	conserved hypotet	322	6	2.2	198	2	S07130	casein B precursor
250	6	2.2	138	1	FXCLEX	flavodoxin - Clost	323	6	2.2	200	2	T10922	3C3.13 protein - S
251	6	2.2	138	2	E75347	hypothetical prote	324	6	2.2	201	2	S50219	H7M1 protein - yea
252	6	2.2	139	2	A71190	hypothetical prote	325	6	2.2	203	2	T51653	myb-related transc
253	6	2.2	140	2	C64324	ribosomal protein	326	6	2.2	204	2	T15295	hypothetical prote
254	6	2.2	141	2	S78354	ribosomal protein	327	6	2.2	205	2	S54040	hypothetical prote
255	6	2.2	141	2	F72158	hypothetical prote	328	6	2.2	206	2	T34979	probable lipoprote
256	6	2.2	142	2	P40195	ribosomal protein	329	6	2.2	206	2	T51656	probable transcrip
257	6	2.2	143	1	PWYBBI	H+-transpoting tw	330	6	2.2	207	2	C97221	hypothetical prote
258	6	2.2	143	2	H84326	hypothetical prote	331	6	2.2	208	2	T15790	hypothetical prote
259	6	2.2	145	2	AD1093	hypothetical prote	332	6	2.2	210	2	B82522	hypothetical prote
260	6	2.2	147	2	B97356	ribosomal protein	333	6	2.2	210	2	A87331	hypothetical prote
261	6	2.2	147	2	H70630	hypothetical prote	334	6	2.2	210	2	A72301	hypothetical prote
262	6	2.2	147	2	B27893	hypothetical prote	335	6	2.2	211	2	AH3232	autoinducer synthe
263	6	2.2	148	2	B70339	conserved hypotet	336	6	2.2	212	2	T51646	probable transcrip
264	6	2.2	150	2	B86646	deoxyuridine 5'-tr	337	6	2.2	214	2	E70439	hypothetical prote
265	6	2.2	150	2	B34380	tropoin C gamma -	338	6	2.2	215	2	H97282	adenylate kinase I
266	6	2.2	150	2	S18434	tropoin C isoform	339	6	2.2	215	2	A40050	peptidylprolyl iso
267	6	2.2	150	2	S18435	tropoin C isoform	340	6	2.2	215	2	A84008	hemolysin III BH28
268	6	2.2	150	2	B29979	nebulin (clone lam	341	6	2.2	216	2	H82695	DNA-3-methyladenin
269	6	2.2	150	2	F95233	universal stress p	342	6	2.2	217	2	JC2557	metallopeptidase
270	6	2.2	152	1	ZBBPT4	gene 578 protein -	343	6	2.2	217	2	T51671	myb-related transc
271	6	2.2	154	2	F88550	protein ZC94.5 (lm	344	6	2.2	218	2	T02912	probable transcrip
272	6	2.2	154	2	B70222	conserved hypotet	345	6	2.2	218	2	E84335	hypothetical prote
273	6	2.2	154	2	T36349	probable regulator	346	6	2.2	219	2	E72866	hypothetical prote
274	6	2.2	155	2	H84530	hypothetical prote	347	6	2.2	220	2	A95258	competence protein
275	6	2.2	156	2	F70712	hypothetical prote	348	6	2.2	220	2	A99523	hypothetical prote
276	6	2.2	158	2	A38397	tropoin C-1 - gla	349	6	2.2	221	2	G83998	mutants block spor
277	6	2.2	158	2	H82688	hypothetical prote	350	6	2.2	222	2	D72343	conserved hypotet
278	6	2.2	159	2	D90586	hypothetical prote	351	6	2.2	224	1	JQ1522	peptidylprolyl iso
279	6	2.2	160	2	A38431	Myc-binding factor	352	6	2.2	226	2	T11093	H+-transpoting tw
280	6	2.2	161	2	A83269	hypothetical prote	353	6	2.2	226	2	T51670	myb-related transc
281	6	2.2	162	2	B82350	conserved hypotet	354	6	2.2	226	2	H82710	phosphoglycolate p
282	6	2.2	163	2	S40324	hypothetical prote	355	6	2.2	226	2	F68125	conserved hypotet
283	6	2.2	165	2	H84997	hypothetical prote	356	6	2.2	227	2	A59097	hypothetical prote
284	6	2.2	165	2	S52511	hypothetical prote	357	6	2.2	229	2	B83281	histidine transpor
285	6	2.2	166	2	T51637	myb-related transc	358	6	2.2	229	2	T47227	transforming prote
286	6	2.2	166	2	A75522	response regulator	359	6	2.2	230	2	D75518	probable biopolyme
287	6	2.2	166	2	T51643	myb-related transc	360	6	2.2	230	2	T24199	hypothetical prote
288	6	2.2	168	2	B72635	hypothetical prote	361	6	2.2	231	2	F69147	UDP-N-acetylglucos
289	6	2.2	168	2	F86861	hypothetical prote	362	6	2.2	232	2	F70252	hypothetical prote
290	6	2.2	168	2	A70479	hypothetical prote	363	6	2.2	234	2	E96957	HAD superfamily hy
291	6	2.2	170	1	E65006	hypothetical prote	364	6	2.2	235	2	A30509	exotoxin C precurs
292	6	2.2	171	2	T31478	hypothetical prote	365	6	2.2	236	2	D71733	phosphoribosylamin
293	6	2.2	171	2	H98097	conserved hypotet	366	6	2.2	236	2	F97736	hypothetical prote
294	6	2.2	173	2	E69122	conserved hypotet	367	6	2.2	236	2	S14414	proliferating cell
295	6	2.2	173	2	E64540	hypothetical prote	368	6	2.2	237	2	S55614	hypothetical prote
296	6	2.2	177	1	KIVZVV	chymidine kinase (	369	6	2.2	237	2	B97560	hypothetical prote
297	6	2.2	177	2	E72160	chymidine kinase (	370	6	2.2	237	2	AE1545	transcription regu
298	6	2.2	177	2	T28517	chymidine kinase (	371	6	2.2	237	2	AF1187	transcription regu
299	6	2.2	177	2	AB1513	transcription regu	372	6	2.2	240	2	B64063	nasb protein homol
300	6	2.2	177	2	AG1154	a transcription re	373	6	2.2	240	2	E72623	hypothetical prote
301	6	2.2	178	2	C64206	hypothetical prote	374	6	2.2	240	2	T20319	hypothetical prote
302	6	2.2	180	2	E86134	hypothetical prote	375	6	2.2	242	2	S35060	tropomyosin - hydr
303	6	2.2	180	2	B91293	hypothetical prote	376	6	2.2	243	2	A40394	spec-related prote
304	6	2.2	180	2	A84426	hypothetical prote	377	6	2.2	243	2	T23348	hypothetical prote
305	6	2.2	182	2	A72261	hypothetical prote	378	6	2.2	244	2	AE3507	amidotransferase h
306	6	2.2	182	2	C82574	phage-related lyso	379	6	2.2	244	2	T15223	hypothetical prote
307	6	2.2	184	2	T24251	hypothetical prote	380	6	2.2	245	2	G72064	ABC transporter AT
308	6	2.2	185	2	B71954	hypothetical prote	381	6	2.2	245	2	B86558	ABC transporter AT
309	6	2.2	185	2	B84531	hypothetical prote	382	6	2.2	245	2	G90282	hypothetical prote
310	6	2.2	185	2	A37207	germination respon	383	6	2.2	247	1	TMHOBP	hypothetical prote
311	6	2.2	186	2	F82508	hypothetical prote	384	6	2.2	247	2	I67850	tropomyosin beta c
312	6	2.2	186	2	B87290	hypothetical prote	385	6	2.2	247	2	C82891	tropomyosin - rat
313	6	2.2	186	2	AH2450	hypothetical prote	386	6	2.2	248	2	B97223	hemolysin UH436 [1
314	6	2.2	187	2	D64160	hypothetical prote	387	6	2.2	248	2	S07282	probable 3-ketoacy
315	6	2.2	190	2	D97330	multimeric flavodo	388	6	2.2	248	2	S10623	tropomyosin TW30-p
316	6	2.2	190	2	G96747	hypothetical prote	389	6	2.2	248	2	I67849	tropomyosin 4, fib
317	6	2.2	191	2	F84932	hypothetical prote	390	6	2.2	249	2	D64310	tropomyosin - rat
318	6	2.2	193	2	AD2402	hypothetical prote	391	6	2.2	249	2	C96779	nucleotide-binding
319	6	2.2	193	2	AI2408	hypothetical prote	392	6	2.2	250	2	S50565	unknown protein F9
320	6	2.2	196	2	C86696	hypothetical membr	393	6	2.2	252	2	AF1278	hyperosmolarity-re
321	6	2.2	196	2	S06581	finger protein (cl	394	6	2.2	252	2	AF1641	indol-3-glycerol p

395	6	2.2	252	2	B70934	468	6	2.2	301	2	B69593	GRP-binding protei
396	6	2.2	252	2	T25443	469	6	2.2	301	2	S28287	hypothetical prote
397	6	2.2	255	2	AE1734	470	6	2.2	303	2	B83122	conserved hypochet
398	6	2.2	255	2	S37861	471	6	2.2	304	2	G83820	hypothetical protei
399	6	2.2	257	2	E83001	472	6	2.2	306	1	W2M1EB	E2 protein bovin
400	6	2.2	257	2	S40974	473	6	2.2	307	2	AC2364	hypothetical prote
401	6	2.2	259	2	T51679	474	6	2.2	308	2	H82936	hpr serine/threon
402	6	2.2	259	2	G64385	475	6	2.2	309	2	AC1249	hypothetical prote
403	6	2.2	259	2	T22775	476	6	2.2	310	2	D75202	dipeptide abc tran
404	6	2.2	259	2	B87537	477	6	2.2	311	2	H83401	probable transcrip
405	6	2.2	260	2	B64357	478	6	2.2	313	2	JE0328	intellectin - mouse
406	6	2.2	261	2	B69455	479	6	2.2	316	1	LU0712	annexin XII - Hydr
407	6	2.2	261	2	B69053	480	6	2.2	316	2	S73683	MG38 homolog P02
408	6	2.2	261	2	B85439	481	6	2.2	316	2	T32893	hypothetical prote
409	6	2.2	261	2	S74761	482	6	2.2	318	2	T03026	chitinase (EC 3.2.
410	6	2.2	261	2	T31741	483	6	2.2	318	2	A90797	hypothetical prote
411	6	2.2	262	2	E75478	484	6	2.2	318	2	D85657	hypothetical prote
412	6	2.2	262	2	T51633	485	6	2.2	321	2	S57482	chitinase class 1
413	6	2.2	263	2	T09523	486	6	2.2	322	2	PS0019	Ig gamma-2a chain
414	6	2.2	263	2	S14415	487	6	2.2	322	2	AE2391	hypothetical prote
415	6	2.2	263	2	G71494	488	6	2.2	323	2	A93622	NADH dehydrogenase
416	6	2.2	263	2	T23347	489	6	2.2	323	2	A93624	NADH dehydrogenase
417	6	2.2	266	2	G69483	490	6	2.2	323	2	B90571	hypothetical prote
418	6	2.2	266	2	F81687	491	6	2.2	323	2	T31828	hypothetical prote
419	6	2.2	269	2	AF0602	492	6	2.2	324	2	A90616	NADH2 dehydrogenas
420	6	2.2	269	2	T00088	493	6	2.2	324	2	S56694	chitinase (EC 3.2.
421	6	2.2	269	2	H86776	494	6	2.2	324	2	AF1642	hypothetical prote
422	6	2.2	269	2	E96704	495	6	2.2	324	2	AF1642	3-oxoacyl - acyl-ca
423	6	2.2	269	2	D47754	496	6	2.2	325	2	H69829	heterodisulfide re
424	6	2.2	271	2	F64819	497	6	2.2	325	2	C69118	hypothetical prote
425	6	2.2	271	2	G90741	498	6	2.2	326	2	G90189	hypothetical prote
426	6	2.2	271	2	F85591	499	6	2.2	328	2	D84698	hypothetical prote
427	6	2.2	273	2	S48426	500	6	2.2	329	2	H84987	DNA-directed RNA p
428	6	2.2	273	2	AC02934	501	6	2.2	329	2	T18620	hypothetical prote
429	6	2.2	273	2	D47301	502	6	2.2	329	2	S37831	hypothetical prote
430	6	2.2	273	2	T08409	503	6	2.2	330	2	T01016	hypothetical prote
431	6	2.2	274	2	AB2325	504	6	2.2	331	2	AE2666	divalent cation tr
432	6	2.2	275	2	T03032	505	6	2.2	332	2	S46947	ribosomal protein
433	6	2.2	275	2	T51651	506	6	2.2	333	2	E86444	hypothetical prote
434	6	2.2	276	2	F83296	507	6	2.2	334	2	T47983	NAM-like protein -
435	6	2.2	276	2	T51685	508	6	2.2	334	2	T46238	hypothetical prote
436	6	2.2	277	2	C97073	509	6	2.2	335	2	B45511	chitinase (EC 3.2.
437	6	2.2	277	2	B87740	510	6	2.2	335	2	C84018	immunogenic protei
438	6	2.2	279	2	T05275	511	6	2.2	335	2	G81280	holliday junction
439	6	2.2	279	2	H87917	512	6	2.2	335	2	T21503	hypothetical prote
440	6	2.2	281	2	G75430	513	6	2.2	335	2	AG3564	dTDPglucose 4,6-de
441	6	2.2	282	2	T31088	514	6	2.2	338	2	A69278	branched-chain ami
442	6	2.2	283	2	B72207	515	6	2.2	339	2	C83335	probable oxidoredu
443	6	2.2	283	2	H84430	516	6	2.2	339	2	T22571	hypothetical prote
444	6	2.2	284	2	T16719	517	6	2.2	339	2	T49597	hypothetical prote
445	6	2.2	284	2	E83347	518	6	2.2	341	2	D95260	tryptophanyl-L-RNA
446	6	2.2	285	2	A24199	519	6	2.2	341	2	B86633	tryptophan-tRNA 11
447	6	2.2	286	2	S18216	520	6	2.2	341	2	G98125	tryptophan-tRNA 11
448	6	2.2	286	2	S07532	521	6	2.2	342	2	S27902	type II site-speci
449	6	2.2	287	1	B69865	522	6	2.2	343	2	E71952	3-dehydroquininate s
450	6	2.2	287	1	F82380	523	6	2.2	343	2	C64555	3-dehydroquininate s
451	6	2.2	289	2	AC1399	524	6	2.2	343	2	G96533	hypothetical prote
452	6	2.2	290	2	G86155	525	6	2.2	347	2	T45551	NADH2 dehydrogenas
453	6	2.2	291	2	T51668	526	6	2.2	348	2	G71681	hypothetical prote
454	6	2.2	292	2	T10106	527	6	2.2	348	2	F89765	hypothetical prote
455	6	2.2	292	2	AB3340	528	6	2.2	350	2	A71857	zinc-dependent alc
456	6	2.2	292	2	H70677	529	6	2.2	350	2	E82398	dihydroorotase VCA
457	6	2.2	294	2	S52426	530	6	2.2	350	2	B75192	hypothetical prote
458	6	2.2	294	2	AC2919	531	6	2.2	350	2	T05589	hypothetical prote
459	6	2.2	294	2	D97693	532	6	2.2	350	2	T32215	hypothetical prote
460	6	2.2	295	2	H82502	533	6	2.2	351	2	T08446	hypothetical prote
461	6	2.2	295	2	B82841	534	6	2.2	354	2	H71100	hypothetical prote
462	6	2.2	295	2	T43930	535	6	2.2	355	1	GBBP4	beta protein - sat
463	6	2.2	296	2	G90703	536	6	2.2	355	1	G69309	conserved hypochet
464	6	2.2	296	2	B85554	537	6	2.2	356	2	D87490	NADH dehydrogenase
465	6	2.2	296	2	T15973	538	6	2.2	356	2	D86844	3-dehydroquininate s
466	6	2.2	296	2	T05622	539	6	2.2	356	2	AE2729	cytochrome oxidase
467	6	2.2	298	2	T16495	540	6	2.2	357	2	C97744	hypothetical prote

541	6	2.2	357	2	AI2253	hypothetical prote	614	6	2.2	422	2	E84338	isocitrate dehydro
542	6	2.2	358	1	E69988	conserved hypothet	615	6	2.2	422	2	T40003	hypothetical prote
543	6	2.2	359	2	D82945	translacion releas	616	6	2.2	425	2	D64542	apolipoprotein N-a
544	6	2.2	360	2	AI3577	metal-activated py	617	6	2.2	425	2	D71965	apolipoprotein n-a
545	6	2.2	361	2	A86841	hypothetical prote	618	6	2.2	426	2	F64485	hypothetical prote
546	6	2.2	362	2	A84963	probable proteinas	619	6	2.2	427	2	T04869	transforming prote
547	6	2.2	363	2	C97448	hypothetical prote	620	6	2.2	428	2	F96782	hypothetical prote
548	6	2.2	364	2	G71544	hypothetical prote	621	6	2.2	431	2	T00698	methionyl aminopep
549	6	2.2	365	2	AC1609	aminotriptide	622	6	2.2	431	2	T16191	hypothetical prote
550	6	2.2	366	2	SI1921	naringenin 3-dioxy	623	6	2.2	432	2	T33833	hypothetical prote
551	6	2.2	365	2	AG1246	aminotripeptide	624	6	2.2	433	2	UC7678	RING finger protei
552	6	2.2	369	2	T50820	hypothetical prote	625	6	2.2	434	2	S52008	tubulin beta-2 cha
553	6	2.2	370	2	B81945	probable plus ret	626	6	2.2	434	2	A71559	probable transcrip
554	6	2.2	370	2	E81159	twisting motility	627	6	2.2	437	1	TVFVSK	transforming prote
555	6	2.2	371	2	T03025	mitosis-specific c	628	6	2.2	439	1	A48099	transcription fact
556	6	2.2	371	2	T02963	cyclin A-type (c10	629	6	2.2	439	2	AG3342	homoserine dehydro
557	6	2.2	371	2	T01717	hypothetical prote	630	6	2.2	439	2	T57561	transcription fact
558	6	2.2	372	2	G96512	probable cyclin, 2	631	6	2.2	439	2	G97329	anaerobic dicarbox
559	6	2.2	372	2	T09962	cyclin A-type - Ma	632	6	2.2	440	2	I61183	transcription fact
560	6	2.2	372	2	F97510	BA483fl.2.1 (cox1	633	6	2.2	440	2	S60423	hypothetical prote
561	6	2.2	373	2	H86456	probable cyathio	634	6	2.2	441	2	T11715	probable acetylorn
562	6	2.2	373	2	A84659	hypothetical prote	635	6	2.2	441	2	B49837	clathrin-associate
563	6	2.2	375	2	A64183	xylose transport p	636	6	2.2	442	2	T42091	acetylornithine tr
564	6	2.2	377	1	S41019	transcription fact	637	6	2.2	442	2	A71969	probable histidine
565	6	2.2	378	2	H64387	hypothetical prote	638	6	2.2	443	2	T40855	hypothetical prote
566	6	2.2	378	2	C82049	acetylornithine de	639	6	2.2	443	2	H82957	probable glutamine
567	6	2.2	378	2	T51647	myb-related trans	640	6	2.2	445	2	F82166	muf protein VC171
568	6	2.2	378	2	AH0420	phm protein limpo	641	6	2.2	446	2	F64691	ATP-dependent clip
569	6	2.2	380	2	G84318	hypothetical prote	642	6	2.2	446	2	H95072	hypothetical prote
570	6	2.2	383	2	T20572	hypothetical prote	643	6	2.2	446	2	G97940	hypothetical prote
571	6	2.2	384	2	T02968	cyclin A-type (c10	644	6	2.2	447	2	T05977	hypothetical prote
572	6	2.2	385	2	C72230	probable exonuclea	645	6	2.2	447	2	G86778	glucose inhibited
573	6	2.2	385	2	S70984	recf protein - Myc	646	6	2.2	448	2	AE2590	two component resp
574	6	2.2	386	1	S51648	cellular tumor ant	647	6	2.2	450	2	T37702	hypothetical prote
575	6	2.2	386	2	T00589	hookless1-like pro	648	6	2.2	450	2	AG0008	probable membrane
576	6	2.2	387	2	S75981	magnesium/cobalt t	649	6	2.2	450	2	T38465	probable dolichyl-
577	6	2.2	387	2	E65010	multidrug resist	650	6	2.2	451	2	B96774	hypothetical prote
578	6	2.2	387	2	G91034	multidrug resist	651	6	2.2	452	1	S29510	ubiquinol-cytochro
579	6	2.2	387	2	A85879	multidrug resist	652	6	2.2	452	2	S56938	fructose-2,6-bisph
580	6	2.2	388	1	UC5461	cellulase (EC 3.2.	653	6	2.2	452	2	F71826	ATP-dependent prot
581	6	2.2	388	1	S43920	cellulase (EC 3.2.	654	6	2.2	452	2	A75484	conserved hypothet
582	6	2.2	388	2	T46229	hypothetical prote	655	6	2.2	456	2	AC1336	argininosuccinate
583	6	2.2	391	2	AF3633	hypothetical membr	656	6	2.2	456	2	AI1706	argininosuccinate
584	6	2.2	392	1	D72511	probable hexosyltr	657	6	2.2	456	2	F97372	probable two-compo
585	6	2.2	393	2	S47789	xylose transport p	658	6	2.2	456	2	T03935	leucine zipper tra
586	6	2.2	393	2	A86032	xylose transport p	659	6	2.2	458	2	E82444	NAD(P) transhydrog
587	6	2.2	393	2	C91185	xylose transport p	660	6	2.2	458	2	D97799	NAD(P) transhydrog
588	6	2.2	393	2	C69591	aspartate transami	661	6	2.2	458	2	T34574	NADH2 dehydrogenas
589	6	2.2	393	2	D97979	hypothetical prote	662	6	2.2	458	2	T18722	hypothetical prote
590	6	2.2	394	2	S75996	hypothetical prote	663	6	2.2	459	2	S06607	23S rRNA intion 2
591	6	2.2	394	2	AH0490	sugar transport sy	664	6	2.2	460	2	D84486	hypothetical prote
592	6	2.2	394	2	T45672	hypothetical prote	665	6	2.2	461	2	C81137	NAD(P) transhydrog
593	6	2.2	395	2	A82827	conserved hypothet	666	6	2.2	461	2	A81885	probable NAD(P) tr
594	6	2.2	396	2	H86301	hypothetical prote	667	6	2.2	461	2	H71717	3-deoxy-d-manno-oc
595	6	2.2	397	2	S21780	pach protein - pha	668	6	2.2	462	1	DEECXB	NAD(P) transhydrog
596	6	2.2	398	2	F89694	S-adenosylmethoni	669	6	2.2	462	2	A85766	NAD(P) transhydrog
597	6	2.2	399	2	AC1601	hypothetical prote	670	6	2.2	462	2	D90917	NAD(P) transhydrog
598	6	2.2	402	1	TVFVUR	protein-tyrosine k	671	6	2.2	462	2	AI0682	NAD(P) transhydrog
599	6	2.2	402	2	AH0896	prolinate kinase	672	6	2.2	462	2	T36185	hypothetical prote
600	6	2.2	403	2	S37168	replication protei	673	6	2.2	462	2	S76693	hypothetical prote
601	6	2.2	405	2	S57970	repc protein - Rhl	674	6	2.2	464	2	AG0280	NAD(P) transhydrog
602	6	2.2	410	2	E75048	multidrug resist	675	6	2.2	464	2	T39699	NAD(P) transhydrog
603	6	2.2	411	2	T51739	RNA helicase Rhs I	676	6	2.2	465	2	AC0347	glutathione-disulf
604	6	2.2	412	2	S69633	hypothetical prote	677	6	2.2	466	1	YTB5MA	probable membrane
605	6	2.2	413	2	D71112	hypothetical prote	678	6	2.2	466	2	S44878	methylemowycin A r
606	6	2.2	418	2	T47239	oxalate/formate an	679	6	2.2	467	2	B86225	protein T12M4.6 [i
607	6	2.2	418	2	F72711	hypothetical prote	680	6	2.2	467	2	T18744	hypothetical prote
608	6	2.2	420	2	T51974	hypothetical prote	681	6	2.2	468	2	A54926	UDPglucose 6-dehyd
609	6	2.2	420	2	T36532	probable membrane	682	6	2.2	471	2	D64974	hypothetical prote
610	6	2.2	420	2	T46460	hypothetical prote	683	6	2.2	471	2	E90989	probable transport
611	6	2.2	420	2	A71030	hypothetical prote	684	6	2.2	471	2	H85834	hypothetical prote
612	6	2.2	420	2	B97304	probable membrane	685	6	2.2	472	2	F82143	hypothetical prote
613	6	2.2	421	2	AI1112	hypothetical prote	686	6	2.2	473	2	T13433	pyrophosphate-depe

687	6	2.2	474	1	ERAD41	760	6	2.2	540	2	C81900	hypothetical prote
688	6	2.2	474	2	F64119	761	6	2.2	543	2	S53817	thermosome beta ch
689	6	2.2	475	2	S76734	762	6	2.2	544	2	T35245	probable ABC trans
690	6	2.2	477	2	G83048	763	6	2.2	545	2	D69679	polyketide synthase
691	6	2.2	478	2	A82336	764	6	2.2	545	2	T19172	hypothetical prote
692	6	2.2	480	2	S77206	765	6	2.2	546	2	S52053	cytochrome-c oxida
693	6	2.2	480	2	C75614	766	6	2.2	547	2	F82162	paraquat-inducible
694	6	2.2	481	2	A97659	767	6	2.2	548	2	B46101	protein-tyrosine-P
695	6	2.2	482	2	AC2883	768	6	2.2	548	2	B87596	xyloridase/arabino
696	6	2.2	482	2	F97239	769	6	2.2	548	2	I39175	SH2-domain protein
697	6	2.2	486	2	G87246	770	6	2.2	549	1	ODUTMB	cytochrome-c oxida
698	6	2.2	487	2	T45982	771	6	2.2	549	1	F63961	arginyl-tRNA synth
699	6	2.2	487	2	T19237	772	6	2.2	553	1	T02139	calcium-dependent
700	6	2.2	487	2	T05117	773	6	2.2	553	2	T01535	probable cytochrom
701	6	2.2	487	2	T05127	774	6	2.2	554	2	G86146	hypothetical prote
702	6	2.2	488	2	H90423	775	6	2.2	554	2	AD1572	propanediol dehydr
703	6	2.2	488	2	T47747	776	6	2.2	554	2	AI1218	propanediol dehydr
704	6	2.2	489	2	B70514	777	6	2.2	555	2	C96667	unknown protein, 7
705	6	2.2	489	2	UC5755	778	6	2.2	557	2	AB1043	probable membrane
706	6	2.2	491	2	AG3506	779	6	2.2	558	2	S75066	K+-transporting AT
707	6	2.2	492	2	T28025	780	6	2.2	559	2	G90424	long-chain-fatty-a
708	6	2.2	493	2	T33569	781	6	2.2	559	2	S15902	nebulin, skeletal
709	6	2.2	493	2	S76517	782	6	2.2	559	2	A12227	hypothetical prote
710	6	2.2	495	2	D89808	783	6	2.2	559	2	H87265	amg protein, prob
711	6	2.2	498	2	B86436	784	6	2.2	562	2	T42250	polypeptide N-acet
712	6	2.2	498	2	H83793	785	6	2.2	564	2	I53106	gene g1i protein -
713	6	2.2	500	1	KIBYP	786	6	2.2	565	2	T47625	glutathione-disulf
714	6	2.2	500	2	T49388	787	6	2.2	565	2	T04548	hypothetical synth
715	6	2.2	502	2	T45852	788	6	2.2	566	2	E81404	acetoacetate synth
716	6	2.2	504	2	T10558	789	6	2.2	568	2	T45957	hypothetical prote
717	6	2.2	505	2	A55923	790	6	2.2	570	2	A90015	PTS system, lactos
718	6	2.2	507	1	A46713	791	6	2.2	572	2	B28474	phosphotransferase
719	6	2.2	507	2	C96988	792	6	2.2	572	2	T48601	hypothetical prote
720	6	2.2	508	2	B49131	793	6	2.2	573	2	J00135	hypothetical 62.8K
721	6	2.2	508	2	B96658	794	6	2.2	574	2	UC7327	cytochrome P450 en
722	6	2.2	509	2	UC5288	795	6	2.2	575	2	F81801	acetoacetate synth
723	6	2.2	510	2	UC0254	796	6	2.2	575	2	A81067	acetoacetate synth
724	6	2.2	511	1	S60287	797	6	2.2	576	2	AE1759	phosphomannomutase
725	6	2.2	511	2	E64183	798	6	2.2	577	2	A53202	cytoplasmic C-asso
726	6	2.2	513	2	CT7120	799	6	2.2	581	2	F97184	DNA modification m
727	6	2.2	513	2	S69181	800	6	2.2	582	2	D96003	ABC transporter AT
728	6	2.2	513	2	T24210	801	6	2.2	584	2	B90281	hypothetical prote
729	6	2.2	514	2	UC5289	802	6	2.2	585	1	JC1486	neopullulanase (EC
730	6	2.2	514	2	T10938	803	6	2.2	585	2	S46034	hypothetical prote
731	6	2.2	515	1	T06595	804	6	2.2	586	2	T52330	pectinesterase (EC
732	6	2.2	515	2	S67290	805	6	2.2	586	2	C71656	single-stranded-dn
733	6	2.2	516	2	H81123	806	6	2.2	586	2	T21683	hypothetical prote
734	6	2.2	516	2	H71332	807	6	2.2	587	2	D84426	hypothetical prote
735	6	2.2	516	2	S48476	808	6	2.2	587	2	H83748	two-component sens
736	6	2.2	517	2	T46056	809	6	2.2	588	2	T02184	probable pectinest
737	6	2.2	519	2	S71451	810	6	2.2	588	2	S40925	hypothetical prote
738	6	2.2	521	2	CT5320	811	6	2.2	591	2	E71651	hypothetical prote
739	6	2.2	521	2	H69317	812	6	2.2	593	2	S55189	hypothetical prote
740	6	2.2	527	2	A81712	813	6	2.2	595	2	T52327	pectinesterase (EC
741	6	2.2	527	2	UC1267	814	6	2.2	595	2	T02183	probable pectinest
742	6	2.2	530	2	CA7113	815	6	2.2	598	2	F83977	hypothetical prote
743	6	2.2	530	2	T50498	816	6	2.2	600	2	T17436	ATP-binding protei
744	6	2.2	531	1	T02993	817	6	2.2	600	2	AF0233	inner membrane ABC
745	6	2.2	532	2	S27373	818	6	2.2	604	2	A71947	GTP-binding protei
746	6	2.2	532	2	S27372	819	6	2.2	608	2	T22897	hypothetical prote
747	6	2.2	532	2	T16752	820	6	2.2	611	2	H82351	probable hemolysin
748	6	2.2	532	2	T27457	821	6	2.2	615	2	H96620	hypothetical prote
749	6	2.2	533	2	H86282	822	6	2.2	617	2	T42249	polypeptide N-acet
750	6	2.2	533	2	S73755	823	6	2.2	617	2	T25942	hypothetical prote
751	6	2.2	533	2	T40153	824	6	2.2	618	2	T42248	polypeptide N-acet
752	6	2.2	535	2	A46101	825	6	2.2	618	2	AC0884	glutathionylisperm
753	6	2.2	535	2	F90418	826	6	2.2	619	2	E82141	peptidyl-prolyl gl
754	6	2.2	536	2	T27668	827	6	2.2	619	2	G84489	probable TNP2-like
755	6	2.2	537	2	A86444	828	6	2.2	621	2	A82153	PTS system, fructo
756	6	2.2	538	2	T08874	829	6	2.2	622	2	PT0059	phosphinochiric-in
757	6	2.2	538	2	A41533	830	6	2.2	622	2	E69006	glutamate synthase
758	6	2.2	538	2	B84759	831	6	2.2	623	2	A49840	segment polarity p
759	6	2.2	539	2	JH0263	832	6	2.2	633	2	D64222	DNA topoisomerase

833	6	2.2	635	2	S73358	6	2.2	769	2	D96634	hypothetical prote
834	6	2.2	635	2	AC0046	6	2.2	771	2	A54273	glucocorticoid rec
835	6	2.2	635	2	T00602	6	2.2	775	2	F86026	hypothetical prote
836	6	2.2	638	2	T43018	6	2.2	777	1	ORHUGA	glucocorticoid rec
837	6	2.2	639	2	S62567	6	2.2	779	2	G65151	hypothetical 86.0
838	6	2.2	641	2	A71163	6	2.2	785	2	D91180	hypothetical prote
839	6	2.2	641	2	D83046	6	2.2	790	2	T49542	xylin 1,4-beta-xy
840	6	2.2	642	2	H70331	6	2.2	798	2	T33022	hypothetical prote
841	6	2.2	646	2	AD3409	6	2.2	798	2	AI2053	competence protein
842	6	2.2	648	2	H84587	6	2.2	804	1	ISECTB	DNA topoisomerase
843	6	2.2	648	2	T47895	6	2.2	804	1	YPR58	phenylalanine-trna
844	6	2.2	649	2	C32575	6	2.2	804	2	B91208	DNA gyrase subunit
845	6	2.2	650	2	T00081	6	2.2	804	2	D86054	hypothetical prote
846	6	2.2	652	2	C95129	6	2.2	805	2	E71719	vir4 protein prec
847	6	2.2	652	2	H97999	6	2.2	805	2	E97717	vir4 protein prec
848	6	2.2	656	2	B96649	6	2.2	808	2	F84038	phenylalanyl-trna
849	6	2.2	657	1	TVRFMA	6	2.2	809	2	F81312	leucine-trna ligas
850	6	2.2	658	2	H90013	6	2.2	810	2	S67696	probable membrane
851	6	2.2	660	2	A44432	6	2.2	819	2	T05744	hypothetical prote
852	6	2.2	661	2	S19646	6	2.2	819	2	D85440	Cu2+-transporting
853	6	2.2	661	2	T22319	6	2.2	832	2	S56230	hypothetical prote
854	6	2.2	662	2	T46228	6	2.2	836	2	T33222	hypothetical prote
855	6	2.2	666	2	S36218	6	2.2	837	2	A97557	dimethylglycine de
856	6	2.2	666	2	S50452	6	2.2	837	2	AD2777	dimethylglycine de
857	6	2.2	668	2	T05257	6	2.2	840	2	T36829	probable phenylala
858	6	2.2	671	2	PH0268	6	2.2	857	2	S44883	2C262.3 protein -
859	6	2.2	672	2	S46276	6	2.2	859	2	G86242	hypothetical prote
860	6	2.2	672	2	F71424	6	2.2	860	2	T22974	hypothetical prote
861	6	2.2	677	2	T39251	6	2.2	874	2	T30398	late expression fa
862	6	2.2	679	2	H84516	6	2.2	875	2	AG1632	DNA polymerase I
863	6	2.2	681	2	T19429	6	2.2	878	2	T43767	DNA topoisomerase
864	6	2.2	683	2	A39784	6	2.2	881	2	S67026	probable membrane
865	6	2.2	686	2	S48467	6	2.2	883	2	T40128	kinesin-like prote
866	6	2.2	695	2	T07283	6	2.2	885	2	T26310	hypothetical prote
867	6	2.2	697	2	A86402	6	2.2	886	2	S48371	hypothetical prote
868	6	2.2	699	2	D70533	6	2.2	886	2	E75625	hypothetical prote
869	6	2.2	699	2	A54660	6	2.2	886	2	A82825	aconitase hydratase
870	6	2.2	701	2	E70827	6	2.2	890	2	A98017	probable 2-compone
871	6	2.2	705	2	H89900	6	2.2	890	2	G96654	hypothetical prote
872	6	2.2	705	2	S55420	6	2.2	890	2	C85861	probable 2-compone
873	6	2.2	708	2	S70834	6	2.2	890	2	F64991	yojN protein - Esc
874	6	2.2	713	2	B32575	6	2.2	892	2	T50985	related to transcr
875	6	2.2	713	2	S50130	6	2.2	892	2	AG1661	transcription regu
876	6	2.2	714	2	JX0357	6	2.2	892	2	AI1289	DNA gyrase, chain
877	6	2.2	714	2	G65001	6	2.2	894	2	E83221	hypothetical prote
878	6	2.2	714	2	E91026	6	2.2	894	2	C90719	hypothetical prote
879	6	2.2	714	2	F85870	6	2.2	894	2	D85569	histidine kinase k
880	6	2.2	714	2	AH0798	6	2.2	894	2	B42372	probable ligand-ga
881	6	2.2	716	2	H72635	6	2.2	895	2	A84640	sensor protein Kdp
882	6	2.2	717	1	S29923	6	2.2	900	2	D97351	hypothetical prote
883	6	2.2	728	1	TVHUSK	6	2.2	902	2	D71079	hypothetical prote
884	6	2.2	729	2	T23769	6	2.2	902	2	AD0587	sensor protein Kdp
885	6	2.2	733	2	S44876	6	2.2	907	2	AE1352	hypothetical prote
886	6	2.2	733	2	A48128	6	2.2	908	2	AH1722	hypothetical prote
887	6	2.2	734	2	T07544	6	2.2	908	2	AI0327	two-component regu
888	6	2.2	736	2	G01522	6	2.2	920	2	B84640	probable ligand-ga
889	6	2.2	739	2	A83015	6	2.2	922	2	D70066	SNF2 helicase homo
890	6	2.2	742	1	QRHUGB	6	2.2	923	2	A39596	progesterone recep
891	6	2.2	746	2	T24978	6	2.2	923	2	I53280	progesterone recep
892	6	2.2	748	2	T47250	6	2.2	924	2	S34926	hypothetical prote
893	6	2.2	750	2	H70711	6	2.2	925	2	T01384	hypothetical prote
894	6	2.2	750	2	A32575	6	2.2	929	2	A32495	rep-1 protein, for
895	6	2.2	751	1	T49497	6	2.2	930	2	A25923	progesterone recep
896	6	2.2	752	1	S03423	6	2.2	932	2	C70191	penicillin-binding
897	6	2.2	752	1	S64750	6	2.2	933	1	QRHUP	progesterone recep
898	6	2.2	757	1	I50667	6	2.2	940	2	F86502	valine-trna synthet
899	6	2.2	757	1	T16149	6	2.2	940	2	B72120	valine-trna ligase
900	6	2.2	761	2	T09052	6	2.2	946	2	S48255	probable membrane
901	6	2.2	766	2	AC0985	6	2.2	958	2	S62065	alanine-trna ligas
902	6	2.2	766	2	A40258	6	2.2	958	2	A70634	probable mmp1 pro
903	6	2.2	767	2	B84594	6	2.2	960	2	S54461	hypothetical prote
904	6	2.2	768	2	JC6564	6	2.2	963	2	T26022	hypothetical prote
905	6	2.2	769	2	E69404	6	2.2	963	2	T41479	Rad16 nucleotide e

```

979 6 2.2 971 2 D70128 conserved hypothe
980 6 2.2 974 1 A49714 protein-tyrosine k
981 6 2.2 979 1 J00894 p15 tyrosine - Mwc
982 6 2.2 980 2 H90681 probable flagellin
983 6 2.2 980 2 T38632 probable phosphati
984 6 2.2 980 2 D85532 probable structura
985 6 2.2 987 1 J00016 6-phosphotricotin
986 6 2.2 992 2 T38817 hypotheical prote
987 6 2.2 993 2 C55226 cy1M protein - Ent
988 6 2.2 996 2 E98092 cy1M protein, cyto
989 6 2.2 996 2 A95228 bacteriocin format
990 6 2.2 1003 2 B71469 glycine-cRNA ligas
991 6 2.2 1004 2 H70673 probable mbcd proc
992 6 2.2 1006 2 T00050 hypotheical prote
993 6 2.2 1008 2 C84679 hypotheical prote
994 6 2.2 1021 1 S64506 protein kinase BUB
995 6 2.2 1022 1 I39643 RXX-toxin I - Acti
996 6 2.2 1032 2 T18293 guanylate kinase-1
997 6 2.2 1044 2 H97186 glycosyltransferas
998 6 2.2 1055 2 A87364 Ompa-related prote
999 6 2.2 1066 2 T26753 hypotheical prote
1000 6 2.2 1073 2 T06074 hypotheical prote

```

## ALIGNMENTS

```

RESULT 1
B96689
hypothetical protein T27F4.13 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Aug-2002
C:Accession: B96689
R:Meologis, A.; Becker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B96689
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-139 <STO>
A:Cross-references: GB:AE005173; NID:g10092494; PIDN:AA612894.1; GSPDB:GN00141
C:Genetics:
A:Gene: T27F4.13
A:Map position: 1
C:Superfamily: Arabidopsis myb-related protein Y19; myb DNA-binding repeat homology

Query Match 11.7%; Score 32; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 1,3e-26;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 LRLHKLGNRWSLIAGRLPGRTANDVKNYWT 109
Db 75 LRLHKLGNRWSLIAGRLPGRTANDVKNYWT 106

RESULT 2
A96689
hypothetical protein T27F4.12 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A96689
R:Meologis, A.; Becker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000

```

```

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A96689
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-246 <STO>
A:Cross-references: GB:AE005173; NID:g10092494; PIDN:AA612894.1; GSPDB:GN00141
C:Genetics:
A:Gene: T27F4.12
A:Map position: 1

Query Match 11.7%; Score 32; DB 2; Length 246;
Best Local Similarity 100.0%; Pred. No. 2,2e-26;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 LRLHKLGNRWSLIAGRLPGRTANDVKNYWT 109
Db 75 LRLHKLGNRWSLIAGRLPGRTANDVKNYWT 106

```

## RESULT 3

```

T51687
myb-related transcription factor MYB90 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000
C:Accession: T51687
R:Kranz, H.D.; Denekamp, M.; Greco, R.; Jin, H.; Kranz, H.D.; Denekamp, M.; Greco, R.; Ji
Paz-Ares, J.; Weishaar, B.
Plant J. 16, 263-276, 1998
A:Title: Towards functional characterisation of the members of the R2R3-MYB gene from Ar
A:Reference number: Z14349; MUID:9839469; PMID:9839469
A:Accession: T51687
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-249 <KRA>
A:Cross-references: EMBL:AF062915; PIDN:AAC83637.1
A:Experimental source: cultivar Columbia
A:Genetics:
A:Gene: MYB90
C:Superfamily: Arabidopsis 28K leaf-specific myb-related protein; myb DNA-binding repeat
C:Keywords: transcription factor

Query Match 11.7%; Score 32; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 2,3e-26;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 LRLHKLGNRWSLIAGRLPGRTANDVKNYWT 109
Db 75 LRLHKLGNRWSLIAGRLPGRTANDVKNYWT 106

RESULT 4
T51680
myb-related transcription factor MYB75 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: T51680
R:Kranz, H.D.; Denekamp, M.; Greco, R.; Jin, H.; Kranz, H.D.; Denekamp, M.; Greco, R.; Ji
Paz-Ares, J.; Weishaar, B.
Plant J. 16, 263-276, 1998
A:Title: Towards functional characterisation of the members of the R2R3-MYB gene from Ar
A:Reference number: Z14349; MUID:9839469; PMID:9839469
A:Accession: T51680
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-212 <KRA>
A:Cross-references: EMBL:AF062908; PIDN:AAC83630.1
A:Experimental source: cultivar Columbia

```



C:Genetics:  
A:Gene: MYB75  
C:Keywords: transcription factor

Query Match 9.9%; Score 27; DB 2; Length 212;  
Best Local Similarity 100.0%; Pred. No. 5e-21;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 LIGNRWSLIAGRLPGRTANDVKRYMYNT 109  
DB 80 LIGNRWSLIAGRLPGRTANDVKRYMYNT 106

# RESULT 5

probable transcription factor F25P12.92 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C:Accession: B96608

R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nucleur 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.; Sun, H.; Tallon,

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; M01D:21016719; PMID:11130712

A:Accession: B96608

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-246 <STO>

A:Cross-references: GB:AE005173; NID:g9954749; PIDN:AA09100.1; GSPDB:GN00141

C:Genetics:

A:Gene: F25P12.92

A:Map position: 1

C:Superfamily: Arabidopsis 28k leaf-specific myb-related protein; myb DNA-binding repeat

Query Match 9.9%; Score 27; DB 2; Length 248;  
Best Local Similarity 100.0%; Pred. No. 5.8e-21;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 LIGNRWSLIAGRLPGRTANDVKRYMYNT 109  
DB 80 LIGNRWSLIAGRLPGRTANDVKRYMYNT 106

# RESULT 6

myb-related protein - upland cotton

N:Alternate names: MYB-like DNA-binding domain protein  
C:Species: Gossypium hirsutum (upland cotton)

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 19-May-2000

C:Accession: T09744

R:Loguericio, L.L.; Zhang, J.; Wilkins, T.A.

Submitted to the EMBL Data Library, November 1997

A:Description: Structure and expression of six classes of myb-domain genes in allotetrap

A:Reference number: Z16842

A:Accession: T09744

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-193 <LOG>

A:Cross-references: EMBL:AF034131; NID:g2921333; PID:g2921334

A:Experimental source: cultivar Acala S1-2; ovule

C:Genetics:

A:Gene: Cmy-G

C:Superfamily: unassigned myb DNA-binding repeat proteins; myb DNA-binding repeat homolo

Query Match 9.1%; Score 25; DB 2; Length 193;

Best Local Similarity 100.0%; Pred. No. 6.8e-19;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 DLIRLRLKLLGNRWSLIAGRLPGRT 99  
DB 76 DLIRLRLKLLGNRWSLIAGRLPGRT 100

# RESULT 7

myb-related protein - upland cotton

N:Alternate names: MYB-like DNA-binding domain protein  
C:Species: Gossypium hirsutum (upland cotton)

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 19-May-2000

C:Accession: T09743

R:Loguericio, L.L.; Zhang, J.; Wilkins, T.A.

Submitted to the EMBL Data Library, November 1997

A:Description: Structure and expression of six classes of myb-domain genes in allotetrap

A:Reference number: Z16842

A:Accession: T09743

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-198 <LOG>

A:Cross-references: EMBL:AF034130; NID:g2921331; PID:g2921332

A:Experimental source: cultivar Acala S1-2

C:Genetics:

A:Gene: Cmy-D

C:Superfamily: unassigned myb DNA-binding repeat proteins; myb DNA-binding repeat homolo

C:Keywords: DNA binding

F:63-113/Domain: myb DNA-binding repeat homology <MYB>

Query Match 7.7%; Score 21; DB 2; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 RLHKLLGNRWSLIAGRLPGRT 99  
DB 81 RLHKLLGNRWSLIAGRLPGRT 101

# RESULT 8

anthocyanin biosynthesis regulatory protein PL, nonfunctional - maize

N:Alternate names: PL transcription factor  
C:Species: Zea mays (maize)

C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 21-Jan-2000

C:Accession: T01189

R:Cooper, P.S.; Kent, B.; Matz, E.C.; Cone, K.C.

Submitted to the EMBL Data Library, July 1997

A:Description: Pl-987: a nonfunctional maize anthocyanin regulatory gene resulting from 1

A:Reference number: Z14257

A:Accession: T01189

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-256 <COO>

A:Cross-references: EMBL:AF015269; NID:g2343274; PIDN:AA067721.1; PID:g2343275

A:Experimental source: cultivar W22

A>Note: the protein is nonfunctional due to insertion of retrotransposon Magellan 5' to t

C:Genetics:

A:Gene: PL

A:Insertions: 34/1; 77/2

C:Superfamily: anthocyanin biosynthesis regulatory protein; myb DNA-binding repeat homolo

Query Match 7.7%; Score 21; DB 2; Length 256;  
Best Local Similarity 100.0%; Pred. No. 1.9e-14;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 RLHKLLGNRWSLIAGRLPGRT 99  
DB 69 RLHKLLGNRWSLIAGRLPGRT 89

## RESULT 9

T03974  
anthocyanin biosynthesis regulatory protein - maize  
C/Species: Zea mays (maize)  
C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 13-Aug-1999  
C/Accession: T03974  
R/Cone, K.C.; Cocciolone, S.M.; Moehlenkamp, C.A.; Weber, T.; Drummond, B.J.; Tagliani, Plant Cell 5, 1807-1816, 1993  
A/Title: Role of the regulatory gene pl in the photocontrol of maize anthocyanin pigment  
A/Reference number: Z15034; MUID:94138244; PMID:8305873  
A/Accession: T03974  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-264 <CON>  
A/Cross-references: EMBL:L19496; NID:G309571; PID:AAA19819.1; PID:G309572  
A/Experimental source: strain Tx303  
C/Genetics:  
A/Intons: 45/1; 88/2  
C/Function:  
A/Description: transcriptional activator for anthocyanin biosynthesis  
C/Superfamily: anthocyanin biosynthesis regulatory protein; myb DNA-binding repeat homol  
F/9-61/Domain: myb DNA-binding repeat homology <MYB>  
F/62-112/Domain: myb DNA-binding repeat homology <MYB1>

Query Match 7.7%; Score 21; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 2e-14;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 RLHKLGNRWSLIAGRLPGRT 99  
Db 80 RLHKLGNRWSLIAGRLPGRT 100

## RESULT 10

T03972  
anthocyanin biosynthesis regulatory protein pl - maize  
C/Species: Zea mays (maize)  
C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 13-Aug-1999  
C/Accession: T03972  
R/Cone, K.C.; Cocciolone, S.M.; Moehlenkamp, C.A.; Weber, T.; Drummond, B.J.; Tagliani, Plant Cell 5, 1807-1816, 1993  
A/Title: Role of the regulatory gene pl in the photocontrol of maize anthocyanin pigment  
A/Reference number: Z15034; MUID:94138244; PMID:8305873  
A/Accession: T03972  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-266 <CON>  
A/Cross-references: EMBL:L19495; NID:G309569; PID:AAA19821.1; PID:G309570  
A/Experimental source: strain MCclintock  
C/Genetics:  
A/Intons: 45/1; 88/2  
C/Function:  
A/Description: transcriptional activator for anthocyanin biosynthesis  
C/Superfamily: anthocyanin biosynthesis regulatory protein; myb DNA-binding repeat homol  
F/9-61/Domain: myb DNA-binding repeat homology <MYB>  
F/62-112/Domain: myb DNA-binding repeat homology <MYB1>

Query Match 7.7%; Score 21; DB 2; Length 266;  
Best Local Similarity 100.0%; Pred. No. 2e-14;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 RLHKLGNRWSLIAGRLPGRT 99  
Db 80 RLHKLGNRWSLIAGRLPGRT 100

## RESULT 11

T01188  
anthocyanin biosynthesis regulatory protein pl - maize  
N/Alternate names: pl transcription factor  
C/Species: Zea mays (maize)  
C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 13-Aug-1999

C/Accession: T01188  
R/Cooper, P.S.; Kent, B.; Matz, E.C.; Cone, K.C.  
submitted to the EMBL Data Library, July 1997  
A/Description: pl-987: a nonfunctional maize anthocyanin regulatory gene resulting from  
A/Reference number: Z14257

A/Accession: T01188  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-267 <COO>  
A/Cross-references: EMBL:AF015268; NID:G2343272; PID:AB67720.1; PID:G2343273  
A/Experimental source: cultivar W22  
C/Genetics:  
A/Gene: pl  
A/Map position: 6  
A/Intons: 45/1; 88/2  
C/Function:  
A/Description: transcription regulation; stimulates expression of genes involved in anth  
C/Superfamily: anthocyanin biosynthesis regulatory protein; myb DNA-binding repeat homol  
C/Keywords: DNA binding; transcription regulation  
F/9-61/Domain: myb DNA-binding repeat homology <MYB>  
F/62-112/Domain: myb DNA-binding repeat homology <MYB1>

Query Match 7.7%; Score 21; DB 2; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2e-14;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 RLHKLGNRWSLIAGRLPGRT 99  
Db 80 RLHKLGNRWSLIAGRLPGRT 100

## RESULT 12

T03715  
anthocyanin biosynthesis regulatory protein pl-Bh - maize  
C/Species: Zea mays (maize)  
C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 13-Aug-1999  
C/Accession: T03715  
R/Cocciolone, S.M.; Cone, K.C.  
submitted to the EMBL Data Library, April 1993  
A/Description: pl-Bh, an anthocyanin regulatory gene of maize that leads to variegated p  
A/Reference number: Z15025  
A/Accession: T03715  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-271 <COO>  
A/Cross-references: EMBL:L13454; NID:G293899; PID:AAA3492.1; PID:G293900  
C/Genetics:  
A/Intons: 45/1; 88/2  
A/Note: pl-Bh (Blotched)  
C/Function:  
A/Description: transcriptional activator for anthocyanin biosynthesis  
C/Superfamily: anthocyanin biosynthesis regulatory protein; myb DNA-binding repeat homol  
F/9-61/Domain: myb DNA-binding repeat homology <MYB>  
F/62-112/Domain: myb DNA-binding repeat homology <MYB1>

Query Match 7.7%; Score 21; DB 2; Length 271;  
Best Local Similarity 100.0%; Pred. No. 2e-14;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 RLHKLGNRWSLIAGRLPGRT 99  
Db 80 RLHKLGNRWSLIAGRLPGRT 100

## RESULT 13

T51684  
probable transcription factor MYB82 [imported] - Arabidopsis thaliana (fragment)  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
C/Accession: T51684  
R/Kranz, H.D.; Denekamp, M.; Greco, R.; Jin, H.; Kranz, H.D.; Denekamp, M.; Greco, R.; J  
; Paz-Ares, J.; Weishaar, B.  
Plant J. 16, 263-276, 1998

A>Title: Towards functional characterisation of the members of the R2R3-MYB gene from *Ad*  
A/Reference number: Z14349; MUID:9839469; PMID:9839469  
A/Accession: T51684  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-130 <RKA>  
A/Cross-references: EMBL:AF062912; PDB:AA083634.1  
C/Genetics:  
A/Gene: MYB82

Query Match 6.2%; Score 19; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 1.5e-12;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 HKLGNRWSLIAGRLPGRT 99  
DB 11 HKLGNRWSLIAGRLPGRT 29

## RESULT 14

F96734

probable transcription factor F23N20.2 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C/Accession: F96734

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, N.F.; Hughes, B.; Huizart, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,

Ker, M.; Ku, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: AB6141; MUID:21016719; PMID:11130712

A/Accession: F96734

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-195 <STO>

A/Cross-references: GB:AE005173; NID:g6714312; PDB:AAF26005.1; GSPDB:GN00141

C/Genetics:

A/Gene: F23N20.2

A/Map position: 1

C/Superfamily: Arabidopsis myb-related 24.7K protein; myb DNA-binding repeat homology

Query Match 6.2%; Score 17; DB 2; Length 195;

Best Local Similarity 100.0%; Pred. No. 3.2e-10;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 LIGNRWSLIAGRLPGRT 99  
DB 52 LIGNRWSLIAGRLPGRT 68

## RESULT 15

S71287

myb-related 24.7K transcription factor - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 16-Jun-2000

C/Accession: S71287; J05729

R/Kirik, V.; Baumelein, H.

submitted to the EMBL Data Library, December 1995

A/Description: Isolation and chromosomal localization of leaf-specific cDNA encoding for

A/Reference number: S71287

A/Accession: S71287

A/Molecule type: mRNA

A/Residues: 1-213 <KIR>

A/Cross-references: EMBL:Z68157; NID:g1197189; PDB:CAA92280.1; PID:g1197190

R/Kirik, V.; Baumelein, H.

Gene 183, 109-113, 1996

A/Title: A novel leaf-specific myb-related protein with a single binding repeat.

A/Reference number: J05729; MUID:97149286; PMID:8996094

A/Accession: J05729  
A/Molecule type: mRNA  
A/Residues: 1-213 <KIR>  
A/Cross-references: EMBL:Z68157; NID:g1197189; PDB:CAA92280.1; PID:g1197190  
C/Comment: This protein is involved in the control of anthocyanin biosynthesis and the re

A/Gene: Atmyb12  
C/Superfamily: Arabidopsis myb-related 24.7K protein; myb DNA-binding repeat homology  
F/30-80/Domain: myb DNA-binding repeat homology <MYB2>  
F/54-80/Region: helix-turn-helix

Query Match 6.2%; Score 17; DB 1; Length 213;  
Best Local Similarity 100.0%; Pred. No. 3.4e-10;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 LIGNRWSLIAGRLPGRT 99  
DB 52 LIGNRWSLIAGRLPGRT 68

Search completed: January 29, 2004, 20:21:38  
Job time : 45 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 20:13:57 ; Search time 18 Seconds  
(Without alignments)  
715.851 Million cell updates/sec

Title: US-10-033-190-2

Perfect score: 274  
Sequence: 1 MNSTMSLSGVKSGSWTDE.....QQGVQNDPSAIDLNLLD 274

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	6.2	252	1 MYBD_MALZE	P23592 zea mays (m
2	17	6.2	258	1 TT2_ARATH	Q9F1A2 arabis thaliana
3	17	6.2	273	1 MYBC_MALZE	P10290 zea mays (m
4	14	5.1	228	1 GL1_ARATH	P27900 arabidopsis
5	12	4.4	267	1 MYB1_HORVU	P20026 hordeum vul
6	9	3.3	399	1 MYBP_MALZE	P27898 zea mays (m
7	8	2.9	1597	1 M3K4_MOUSE	O08648 mus musculu
8	8	2.9	1607	1 M3K4_HUMAN	Q96474 homo sapien
9	7	2.6	208	1 NARP_HUMAN	P48455 haemophilus
10	7	2.6	215	1 NARP_ECOLI	P18022 escherichia
11	7	2.6	224	1 PSB2_CRYNE	Q00826 cryptococcu
12	7	2.6	255	1 MYB3_MALZE	P20025 zea mays (m
13	7	2.6	302	1 MYB3_HORVU	P20027 hordeum vul
14	7	2.6	315	1 YAAC_BACSU	P37526 bacillus su
15	7	2.6	333	1 MDHC_HUMAN	P40925 homo sapien
16	7	2.6	340	1 MYB1_MALZE	P20024 zea mays (m
17	7	2.6	360	1 CHL1_MESVI	Q8MUC3 mesostigma
18	7	2.6	411	1 VE2_BP2	P11299 bovine pap
19	7	2.6	412	1 SPN3_SCHPO	P48008 schistosach
20	7	2.6	421	1 MYB2_PHYPA	P80073 physcomitre
21	7	2.6	475	1 MUKF_HARIN	P40503 haemophilus
22	7	2.6	502	1 KPYK_LACTA	O07637 lactococcus
23	7	2.6	524	1 HME2_SCHMA	Q26621 schistosoma
24	7	2.6	531	1 TMA2_ARATH	Q84294 arabidopsis
25	7	2.6	647	1 GIDA_RALSO	Q84655 ralsstonia s
26	7	2.6	754	1 YAJ3_SCHPO	Q09903 schistosach
27	7	2.6	1013	1 A60D_DROME	P91927 drosophila
28	7	2.6	1451	1 EM30_ARATH	Q42510 arabidopsis
29	6	2.2	74	1 INV3_YEAST	P10595 saccharomyc
30	6	2.2	74	1 INV5_YEAST	P10597 saccharomyc
31	6	2.2	88	1 YM30_ARCFU	Q28053 archaeoglob
32	6	2.2	96	1 INV7_YEAST	P07635 saccharomyc
33	6	2.2	96	1 Y273_METKA	Q8LYM3 methanopyru

107	6	2.2	250	1	GPP2_YEAST	P40106	saccharomyc	180	6	2.2	377	1	Y704_METUA	Q58115	methanococ
108	6	2.2	252	1	TRPC_LISIN	Q92b79	listeria in	181	6	2.2	378	1	ARGE_VIBCH	Q9knt5	vibrio chol
109	6	2.2	252	1	TRPC_LISMO	Q8v664	listeria mo	182	6	2.2	382	1	P53_SHEEP	P51664	ovis aries
110	6	2.2	256	1	NPUL_YEAST	P32880	saccharomyc	183	6	2.2	385	1	BECE_MYCTU	O59586	mycobacteri
111	6	2.2	259	1	Y683_METUA	Q58096	methanococ	184	6	2.2	386	1	P53_BOVIN	Q29628	bos taurus
112	6	2.2	260	1	Y538_METUA	O57958	methanococ	185	6	2.2	387	1	EMRK_ECOLI	P52599	escherichia
113	6	2.2	263	1	PSL3_HUMAN	Q95633	homo sapien	186	6	2.2	388	1	GUN3_HUMIN	Q12624	humicola in
114	6	2.2	263	1	PCNA_BRANA	Q43134	brassica na	187	6	2.2	391	1	P53_MARMO	Q36006	marmota mon
115	6	2.2	263	1	PCNA_ORYSA	P17070	oryza sativ	188	6	2.2	392	1	GAT4_XENLA	Q91607	xenopus lae
116	6	2.2	264	1	PCNA_TOBAC	O82797	nicotiana t	189	6	2.2	393	1	AAT1_BACSU	P33001	bacillus su
117	6	2.2	264	1	RPOD_SULTO	Q96y40	sulfolobus	190	6	2.2	393	1	XYLH_ECOLI	P37389	escherichia
118	6	2.2	266	1	Y583_CHLMD	Q9pk84	chlamydia m	191	6	2.2	396	1	PYRE_SORNA	P78748	soraria ma
119	6	2.2	269	1	HIS9_LACIA	Q02180	lactococcus	192	6	2.2	397	1	METK_STRAW	P50307	staphylococ
120	6	2.2	269	1	SET_DROME	P33997	drosophila	193	6	2.2	397	1	METK_STRAW	O8mz92	staphylococ
121	6	2.2	271	1	VBIV_ECOLI	P75732	escherichia	194	6	2.2	397	1	PACA_BPPI	P28813	bacteriopho
122	6	2.2	281	1	PANC_DEIRA	Q9rv66	deinococcus	195	6	2.2	397	1	PPD_STRVR	O86938	streptomyce
123	6	2.2	284	1	TPM3_HUMAN	P06753	homo sapien	196	6	2.2	398	1	METK_STRAM	Q99c79	staphylococ
124	6	2.2	284	1	TPM3_MOUSE	P21107	mus muscicu	197	6	2.2	399	1	BMBB_MOUSE	P55105	mus muscicu
125	6	2.2	286	1	FDHD_WOISU	P28181	wollinella s	198	6	2.2	401	1	PPD_STRHY	P55105	mus muscicu
126	6	2.2	286	1	YKUE_BACSU	P22311	aciara copr	199	6	2.2	402	1	KROS_AVISU	Q54271	streptomyce
127	6	2.2	287	1	YKUE_BACSU	Q34870	bacillus su	200	6	2.2	402	1	TDOD_SALTI	P00529	avian sarco
128	6	2.2	292	1	Y251_MYCTU	P71850	mycobacteri	201	6	2.2	402	1	TDOD_SALTY	O06961	salmonella
129	6	2.2	301	1	ERA_BACSU	P2182	bacillus su	202	6	2.2	410	1	VE2_BPYI	P03122	bovine papi
130	6	2.2	301	1	YLD3_CAEBL	Q03568	caenorhabdi	203	6	2.2	417	1	OXLT_OXAFO	Q51330	oxalobacter
131	6	2.2	304	1	ERA_BACHD	Q9kd52	bacillus ha	204	6	2.2	423	1	GCH2_CORAM	O24752	corynebacte
132	6	2.2	304	1	THU4_AGRVI	P52669	agrobacteri	205	6	2.2	433	1	RN26_HUMAN	Q9by78	homo sapien
133	6	2.2	307	1	THU4_MOUSE	Q9j340	mus muscicu	206	6	2.2	437	1	SKI_AVIES	P17863	avian eryth
134	6	2.2	307	1	NIH2_MASLA	Q47921	mastigoclad	207	6	2.2	440	1	ENG4_FUSNN	O8rgv7	fusobacteri
135	6	2.2	308	1	HPRK_UREPA	Q9pr59	ureaplasma	208	6	2.2	440	1	GAT4_MOUSE	O08369	mus muscicu
136	6	2.2	313	1	THAB_HUMAN	Q96ek4	homo sapien	209	6	2.2	440	1	GAT4_RAT	P46152	rattus norv
137	6	2.2	314	1	CHIB_VITVI	P51613	vitis vinif	210	6	2.2	441	1	YGR4_YEAST	P53107	saccharomyc
138	6	2.2	314	1	P53_SPEBE	O64662	spermophilu	211	6	2.2	441	1	AP50_CAEBL	P35603	caenorhabdi
139	6	2.2	316	1	ANXC_HYDAT	P26256	hydra atten	212	6	2.2	441	1	ARGD_SCHPO	O74548	schizosacch
140	6	2.2	316	1	YEB5_MYCPN	P75300	mycoplasma	213	6	2.2	442	1	GAT4_HUMAN	P43694	homo sapien
141	6	2.2	316	1	YIS1_PYRFU	Q8t244	pyrococcus	214	6	2.2	446	1	CLPX_HELPY	O25926	helicobacte
142	6	2.2	318	1	Y211_AQUAE	O66405	aquifex aeo	215	6	2.2	447	1	GID_LACIA	Q9cg74	lactococcus
143	6	2.2	322	1	CHIT_ARATH	P19171	arabidopsis	216	6	2.2	449	1	MYB1_PHYPA	P80074	physcomitri
144	6	2.2	322	1	GCA_RAT	P07670	rattus norv	217	6	2.2	450	1	OSTA_SCHPO	Q10176	schizosacch
145	6	2.2	323	1	SV42_ARATH	Q9w470	arabidopsis	218	6	2.2	452	1	CLPX_HELPJ	Q9z318	helicobacte
146	6	2.2	324	1	CHI2_PEA	P21226	pisum sativ	219	6	2.2	452	1	UCR2_YEAST	P32650	saccharomyc
147	6	2.2	327	1	MOXR_RAT	Q9e588	rattus norv	220	6	2.2	452	1	UCR2_RAT	P32551	rattus norv
148	6	2.2	329	1	AOPC_PENCL	Q40784	penicillium	221	6	2.2	453	1	UCR2_MOUSE	Q9db77	mus muscicu
149	6	2.2	329	1	RPOA_BUCAI	P57666	buchnera ap	222	6	2.2	456	1	ARLY_LISIN	Q92988	listeria in
150	6	2.2	329	1	SRA2_CAEBL	O09204	caenorhabdi	223	6	2.2	456	1	ARLY_LISMO	O8y511	listeria mo
151	6	2.2	329	1	YCLC_PSRPU	P47100	pseudomonas	224	6	2.2	456	1	TG21_TOBAC	O24160	nicotiana t
152	6	2.2	329	1	YKX8_YEAST	P36104	saccharomyc	225	6	2.2	458	1	NUON_RICCN	Q92b55	ricicetaria
153	6	2.2	330	1	NDP4_MOUSE	O09105	mus muscicu	226	6	2.2	461	1	KDTR_RICPR	Q9ze58	ricicetaria
154	6	2.2	335	1	FEN_THEAC	Q9pnc7	campylobact	227	6	2.2	462	1	PNTB_ECOLI	P07002	escherichia
155	6	2.2	336	1	SOHB_BUCAI	P57370	arabidopsis	228	6	2.2	464	1	GSHR_SCHPO	P78965	schizosacch
156	6	2.2	336	1	SOHB_BUCAI	P57370	arabidopsis	229	6	2.2	464	1	SMAS_BRARE	Q9w7e7	brachydanio
157	6	2.2	337	1	RLAO_SUITO	O97312	sulfolobus	230	6	2.2	466	1	MMR_BACSU	O00538	bacillus su
158	6	2.2	340	1	SYW_STRPJ	O8ny42	streptococ	231	6	2.2	467	1	CBPA_DICDI	P35085	dictyosteli
159	6	2.2	340	1	SYW_STRPJ	Q99xn4	streptococ	232	6	2.2	468	1	MURC_FUSNN	Q8nd42	escherichia
160	6	2.2	341	1	SYW_LACIA	Q9cjd1	lactococcus	233	6	2.2	471	1	YEG8_ECOLI	P36554	escherichia
161	6	2.2	341	1	SYW_STRPN	Q97n42	streptococ	234	6	2.2	474	1	DNB2_ADE41	P11807	human adeno
162	6	2.2	342	1	GLK_RALSO	P58617	raistocoria s	235	6	2.2	474	1	PNTB_HAEIN	P43010	haemophilus
163	6	2.2	342	1	PYRC_VIBCH	Q8k134	vibrio chol	236	6	2.2	481	1	ARLY_METWP	O74026	methanococ
164	6	2.2	342	1	PYRC_VIBYU	Q8d359	vibrio vuln	237	6	2.2	485	1	TPH2_RAT	O8c9n9	rattus norv
165	6	2.2	342	1	TYC2_CHVPI	P31117	paramecium	238	6	2.2	488	1	FLID_XENNE	O56877	xenorhabdus
166	6	2.2	343	1	AROB_HELPJ	Q9cm42	helicobacte	239	6	2.2	489	1	TPH2_MOUSE	Q8c9y2	mus muscicu
167	6	2.2	343	1	AROB_HELPJ	P56081	helicobacte	240	6	2.2	489	1	ANSP_MYCTU	Q33261	mycobacteri
168	6	2.2	348	1	SKI_MOUSE	O60698	mus muscicu	241	6	2.2	490	1	HDVD_CLOAM	P55792	clostridium
169	6	2.2	350	1	PAN2_SCHPO	Q9hdus	echinosacch	242	6	2.2	490	1	TPH2_HUMAN	O8iwn9	homo sapien
170	6	2.2	351	1	TMO2_HUMAN	Q9nzt1	homo sapien	243	6	2.2	493	1	XYLH_HAEIN	P44401	haemophilus
171	6	2.2	355	1	LEU3_THETN	Q8rdk0	thermoanaer	244	6	2.2	494	1	UGDH_BOVIN	P12378	bos taurus
172	6	2.2	355	1	VBET_BPPI	P13057	bacteriopho	245	6	2.2	500	1	KPY1_YEAST	O60701	homo sapien
173	6	2.2	356	1	AROB_LACIA	O8ee88	lactococcus	246	6	2.2	500	1	KPY1_YEAST	P00549	saccharomyc
174	6	2.2	359	1	THII_BACSU	O34595	bacillus su	247	6	2.2	505	1	SMI1_YEAST	P32566	saccharomyc
175	6	2.2	365	1	FLIH_DIACA	O93p90	ureaplasma	248	6	2.2	506	1	SYG_DEIRA	O9r8r5	deinococcus
176	6	2.2	367	1	FEK_PARRP	O05964	dianthus ca	249	6	2.2	507	1	EPOR_RAT	O07303	rattus norv
177	6	2.2	375	1	XYIH_HAEIN	O00869	paramecium	250	6	2.2	507	1	KROS_CHICK	P08991	gallus gall
178	6	2.2	375	1	XYIH_HAEIN	P45045	haemophilus	251	6	2.2	510	1	HGV2_HALRO	O02508	halocynthia
179	6	2.2	377	1	TBX9_CAEBL	Q22289	caenorhabdi	252	6	2.2	511	1	G6PD_SOLTU	P37830	solanum tub

253	6	2.2	513	1	PDI_MAIZE	P52388	zea mays (m	326	6	2.2	717	1	SKI_XENIA	Q02225	xenopus lae
254	6	2.2	513	1	SHS1_MOUSE	P97797	m proteina-t	327	6	2.2	723	1	S21C_MOUSE	Q08078	mus musculus
255	6	2.2	516	1	Y369_TREPA	O83384	treponema p	328	6	2.2	723	1	S21C_RAT	Q09001	rattus norv
256	6	2.2	516	1	Y1J7_YEAST	P40492	saccharomyc	329	6	2.2	728	1	SKI_HUMAN	P12755	homo sapien
257	6	2.2	525	1	YXAK_MOUSE	O21973	caenorhabdi	330	6	2.2	733	1	ENG_DROME	Q2412	drosophila
258	6	2.2	526	1	C11A_MOUSE	Q9q262	mus musculu	331	6	2.2	734	1	PSAB_PINTH	P41640	pinus thunb
259	6	2.2	527	1	MALQ_MOUSE	Q9pkur	chlamydia m	332	6	2.2	736	1	XMS1_DROME	Q93v08	drosophila
260	6	2.2	530	1	UDBE_RABIT	P36513	oryctolagus	333	6	2.2	738	1	S521_RAT	Q99v02	rattus norv
261	6	2.2	532	1	INVA_YEAST	P10594	saccharomyc	334	6	2.2	748	1	C184_NEUCR	O42637	rattus norv
262	6	2.2	532	1	INVA_YEAST	P10594	saccharomyc	335	6	2.2	750	1	APB2_RAT	O35431	rattus norv
263	6	2.2	532	1	YRBS_MOUSE	O09346	caenorhabdi	336	6	2.2	750	1	MUTB_MYCTU	P71774	mycobacteri
264	6	2.2	533	1	YRBS_MOUSE	O09346	caenorhabdi	337	6	2.2	750	1	SKI_CHICK	P49140	gallus gall
265	6	2.2	533	1	YRBS_MOUSE	O09346	caenorhabdi	338	6	2.2	751	1	MYBA_MOUSE	P51960	mus musculu
266	6	2.2	536	1	PPCK_BRUSU	O8fey01	bruceella su	339	6	2.2	752	1	DRS1_YEAST	P32922	saccharomyc
267	6	2.2	536	1	PPCK_BRUSU	O8fey01	bruceella su	340	6	2.2	752	1	MYBA_HUMAN	P10243	homo sapien
268	6	2.2	537	1	CNE1_HUMAN	P09829	homo sapien	341	6	2.2	757	1	GSH1_PASMT	O9cm00	pasteurella
269	6	2.2	538	1	RETO_BSCCA	P30986	eschscholzi	342	6	2.2	757	1	MYBA_CHICK	P52550	gallus gall
270	6	2.2	542	1	KPYK_YARLI	P30614	yarrowia li	343	6	2.2	757	1	YQF2_MOUSE	Q09314	caenorhabdi
271	6	2.2	542	1	V155_FOMPV	O9j587	fowlipox vir	344	6	2.2	764	1	METE_SOLSC	O42662	solenostemo
272	6	2.2	543	1	SYFB_TREAC	P57694	thermoplasm	345	6	2.2	765	1	METE_ARATH	O50008	arabidopsis
273	6	2.2	543	1	THSB_TREAC	P48425	thermoplasm	346	6	2.2	766	1	BCSB_SALTY	O82290	salmonella
274	6	2.2	549	1	COX1_TRYVB	P04371	trypanosoma	347	6	2.2	766	1	BCSB_SALTY	O82290	salmonella
275	6	2.2	549	1	SYR_ARCTU	O29368	archaeoglob	348	6	2.2	766	1	GAP1_SCHPO	P33277	schizosacch
276	6	2.2	553	1	C862_ARATH	O23066	arabidopsis	349	6	2.2	771	1	GCR_CAVPO	P49115	cavia porce
277	6	2.2	558	1	ATKA_SYNY3	P73866	synchocyst	350	6	2.2	773	1	YOD3_MOUSE	P34595	caenorhabdi
278	6	2.2	555	1	GSHC_ARATH	P42770	arabidopsis	351	6	2.2	777	1	GCR_AOTNA	P79686	aotus nancy
279	6	2.2	572	1	PTLB_STAAT	P11662	staphylococ	352	6	2.2	777	1	GCR_HUMAN	P04150	homo sapien
280	6	2.2	574	1	C4GF_DROME	O9vyy4	drosophila	353	6	2.2	777	1	GCR_SAGOE	P79269	saguinus oe
281	6	2.2	575	1	COLL_DROME	P56721	drosophila	354	6	2.2	777	1	GCR_SALTB	O13186	salintri bol
282	6	2.2	579	1	COE2_BHARE	O93375	brachydanio	355	6	2.2	778	1	GCR_SALSC	O46567	salintri bol
283	6	2.2	585	1	NEPU_THEVU	O08751	thermoactin	356	6	2.2	779	1	BCSB_ECOLI	P37652	eschcherichia
284	6	2.2	585	1	YB13_YEAST	P38289	saccharomyc	357	6	2.2	784	1	TLR2_HUMAN	O60603	homo sapien
285	6	2.2	588	1	Y06H_MOUSE	P34602	caenorhabdi	358	6	2.2	784	1	TLR2_MACFA	O95m53	macaca fasc
286	6	2.2	593	1	MP10_YEAST	P47083	saccharomyc	359	6	2.2	785	1	BCSB_ECO57	O8x518	eschcherichia
287	6	2.2	594	1	GRBA_HUMAN	O13322	homo sapien	360	6	2.2	794	1	SYFB_TREIN	O813c7	thermoanaer
288	6	2.2	604	1	LEPA_HELPU7	O92m03	helicobacte	361	6	2.2	794	1	TRK1_LYMSH	O76997	lymaea sta
289	6	2.2	609	1	VEI_HPV37	O80902	human papil	362	6	2.2	797	1	SHK3_HUMAN	Q9byb0	homo sapien
290	6	2.2	615	1	CNK_RAT	O9r011	rattus norv	363	6	2.2	799	1	CN3B_MOUSE	O61409	mus musculu
291	6	2.2	623	1	DSH_DROME	P51140	drosophila	364	6	2.2	803	1	GVRB_ECOLI	P06682	eschcherichia
292	6	2.2	623	1	GIDA_RHIME	O92kx2	rhizobium m	365	6	2.2	804	1	SYFB_BACSU	P17922	baclillus su
293	6	2.2	633	1	GIDA_SYNY3	O8dwe1	streplococ	366	6	2.2	808	1	SYFB_BACHD	Q9x896	baclillus ha
294	6	2.2	633	1	PARE_MYCCE	P47445	mycoplasma	367	6	2.2	809	1	SYL_CAMEE	Q9puk3	campylobact
295	6	2.2	635	1	MUTL_YERPE	O8ziw4	yersinia pe	368	6	2.2	819	1	AHMI_ARATH	O9m4h5	arabidopsis
296	6	2.2	635	1	PARE_MYCPN	P78016	mycoplasma	369	6	2.2	825	1	GVRB_BUCAP	O8kew2	buchnera ap
297	6	2.2	638	1	CC45_SCHPO	O74113	schizosacch	370	6	2.2	830	1	GCI2_MOUSE	O9jnd2	mus musculu
298	6	2.2	639	1	CWF0_SCHPO	O09909	schizosacch	371	6	2.2	832	1	YFC4_YEAST	P43572	saccharomyc
299	6	2.2	641	1	Y252_DROME	O9w4c5	drosophila	372	6	2.2	837	1	GCL2_HUMAN	Q92830	homo sapien
300	6	2.2	656	1	SL54_MOUSE	O9ec37	mus musculu	373	6	2.2	840	1	SYFB_STRCO	O88054	streplococ
301	6	2.2	657	1	LONI_TREAC	O9hj89	thermoplasm	374	6	2.2	862	1	TASY_TAXBA	O93jx3	taxus bacca
302	6	2.2	657	1	MYB_DROME	P04197	drosophila	375	6	2.2	862	1	TASY_TAXHR	O41594	taxus brevi
303	6	2.2	660	1	SL54_PIG	P31636	sus scrofa	376	6	2.2	862	1	TASY_TAXCH	O9fc137	taxus chine
304	6	2.2	661	1	DYIN_DROME	O24246	drosophila	377	6	2.2	862	1	TRP7_HUMAN	O9hcx4	homo sapien
305	6	2.2	661	1	PSO2_YEAST	P30620	saccharomyc	378	6	2.2	862	1	TRP7_MOUSE	O9wvc5	mus musculu
306	6	2.2	666	1	HNF4_DROME	P49866	drosophila	379	6	2.2	881	1	ARPB_YEAST	Q11286	saccharomyc
307	6	2.2	666	1	YEA7_YEAST	P40002	saccharomyc	380	6	2.2	883	1	KLP5_SCHPO	O14343	schizosacch
308	6	2.2	672	1	ACSA_PHYBL	O01576	phycomyces	381	6	2.2	886	1	YIH3_YEAST	P40511	saccharomyc
309	6	2.2	676	1	SS21_HUMAN	O96m07	homo sapien	382	6	2.2	890	1	YOJN_ECOLI	P39838	eschcherichia
310	6	2.2	680	1	APB2_MOUSE	P98084	mus musculu	383	6	2.2	894	1	KDPO_ECOLI	P21865	eschcherichia
311	6	2.2	683	1	APCE_SYNP6	P28035	synchococ	384	6	2.2	896	1	NAHS_HUMAN	Q14940	homo sapien
312	6	2.2	686	1	YIK5_YEAST	P40485	saccharomyc	385	6	2.2	900	1	NAHS_RAT	O92x02	rattus norv
313	6	2.2	695	1	YCX7_CHUVU	O20159	chlorella v	386	6	2.2	908	1	KDPO_CLOAB	P94608	clostridium
314	6	2.2	699	1	ECM2_HUMAN	O94769	homo sapien	387	6	2.2	902	1	NRP1_WOLSU	O981e4	wolfinella s
315	6	2.2	699	1	SRCH_HUMAN	P23327	homo sapien	388	6	2.2	913	1	A2S3_RAT	O8rzh7	rattus norv
316	6	2.2	700	1	MR11_CHICK	O91am7	gallus gall	389	6	2.2	919	1	SYNP_HUMAN	O9ygh5	homo sapien
317	6	2.2	701	1	EF2_MYCTU	O53790	mycobacteri	390	6	2.2	922	1	GYRA_AERSA	P48365	aeromonas s
318	6	2.2	705	1	IF2_STAMM	O99uk3	staphylococ	391	6	2.2	923	1	PRGR_MOUSE	Q00175	mus musculu
319	6	2.2	705	1	YWFJ_BACSU	O8mz1	staphylococ	392	6	2.2	923	1	PRGR_RAT	O63449	rattus norv
320	6	2.2	708	1	RPSD_MYXXA	P45866	baclillus su	393	6	2.2	924	1	YB53_YEAST	P38308	saccharomyc
321	6	2.2	711	1	GCR_PIG	P17531	myxococcus	394	6	2.2	925	1	NRP2_RAT	O35276	rattus norv
322	6	2.2	711	1	MR11_XENIA	O9nlu3	xenopus scrofa	395	6	2.2	930	1	PRGR_RABIT	P06186	oryctolagus
323	6	2.2	712	1	NUOL_RHOCA	O9w6k1	xenopus lae	396	6	2.2	931	1	NRP2_HUMAN	O60462	homo sapien
324	6	2.2	712	1	NUOL_RHOCA	P50939	rhodobacter	397	6	2.2	931	1	NRP2_MOUSE	O35375	mus musculu
325	6	2.2	713	1	PTA_ECOLI	P39184	eschcherichia	398	6	2.2	933	1	PRGR_HUMAN	P06401	homo sapien

399	6	2.2	940	1	SYV_CHLBN	Q92967	chlamydia p	472	5	1.8	41	1	URI_CATCO	P01145	catostomus
400	6	2.2	946	1	YBT6_YEAST	P38250	saccharomyc	473	5	1.8	43	1	MTA_COLVI	P27086	colinus vir
401	6	2.2	958	1	HIG_DROME	O09101	drosophila	474	5	1.8	43	1	MTB_COLVI	P27087	colinus vir
402	6	2.2	958	1	MML1_MYCTU	P95211	mycobacteri	475	5	1.8	44	1	RS7_SALSA	P47039	salmo galar
403	6	2.2	958	1	SYAC_YEAST	P40825	saccharomyc	476	5	1.8	50	1	Y476_HAEIN	P44002	haemophilus
404	6	2.2	960	1	YMX6_YEAST	Q04279	saccharomyc	477	5	1.8	54	1	RLJ3_THEAC	Q9h1c7	thermophilus
405	6	2.2	971	1	Y228_BORBU	O51246	borrelia bu	478	5	1.8	54	1	RLJ3_THEVO	Q974b6	thermophilus
406	6	2.2	979	1	P115_MYCHR	P16861	mycoplasma	479	5	1.8	55	1	YB69_SYNY3	P74805	synchocyst
407	6	2.2	987	1	K6P1_YEAST	O46371	chlamydia t	480	5	1.8	56	1	YPI8_CLOPE	P18018	clostridium
408	6	2.2	1003	1	SYG_CHLNR	P41695	saccharomyc	481	5	1.8	57	1	RLJ3_PASMU	P57912	pasteurella
409	6	2.2	1021	1	BUB1_YEAST	P51628	actinobacil	482	5	1.8	59	1	MTPE_SUITO	O8y4y6	anaeana sp
410	6	2.2	1023	1	RT11_ACTPL	P55128	actinobacil	483	5	1.8	61	1	PPSA_ERWHE	P23059	autolobus
411	6	2.2	1023	1	RT12_ACTPL	P55129	actinobacil	484	5	1.8	61	1	MT4_CANFA	O54457	erwinia her
412	6	2.2	1062	1	YAI13_HUMAN	Q9216	homo sapien	485	5	1.8	62	1	MT4_HUMAN	Q97415	canis famli
413	6	2.2	1070	1	CARB_METAC	O8y216	methanosarc	486	5	1.8	62	1	MT4_MOUSE	P47945	homo sapien
414	6	2.2	1073	1	CARB_METMA	P58944	methanosarc	487	5	1.8	62	1	SECE_SUIISO	P58191	sulfolobus
415	6	2.2	1091	1	MSH3_MOUSE	P13705	mus musculus	488	5	1.8	62	1	MT2_COLLI	P15787	columba liv
416	6	2.2	1095	1	TM83_SCHPO	O74476	schizosacch	489	5	1.8	63	1	MT_CHICK	P09576	gallus gall
417	6	2.2	1102	1	YG49_SCHPO	O60184	schizosacch	490	5	1.8	63	1	YOR4_TTVI	P19279	thermoprote
418	6	2.2	1108	1	CN3B_RAT	P47806	mus musculus	491	5	1.8	63	1	MOR2_BOMMO	P82818	bombyx mori
419	6	2.2	1111	1	GLI1_MOUSE	P58939	corynebacte	492	5	1.8	66	1	MOR2_BOMMO	O96059	bombyx mori
420	6	2.2	1113	1	CARB_CORGL	P13009	escherichia	493	5	1.8	66	1	RL28_TREVO	O972x0	thermoplaem
421	6	2.2	1226	1	METH_ECOLI	Q09441	caenorhabdi	494	5	1.8	66	1	THIS_ECOLI	O32583	escherichia
422	6	2.2	1244	1	YP83_CABEL	P09152	escherichia	495	5	1.8	66	1	YHRI_VARV	P33797	variola vir
423	6	2.2	1246	1	NARG_ECOLI	P47580	mycoplasma	496	5	1.8	66	1	YVBB_VACCC	P20542	vacinia vi
424	6	2.2	1271	1	Y338_MYCCE	P75236	mycoplasma	497	5	1.8	68	1	GBG5_HUMAN	P30650	homo sapien
425	6	2.2	1300	1	Y338_MYCPN	P20241	drosophila	498	5	1.8	68	1	RL29_TREAC	P48506	thermoplaem
426	6	2.2	1302	1	NRG_DROME	P20241	drosophila	499	5	1.8	68	1	UCRH_SOLTU	P48504	solanum tub
427	6	2.2	1317	1	GAP_CABEL	P34288	caenorhabdi	500	5	1.8	69	1	MCHI_ECOLI	O86200	escherichia
428	6	2.2	1341	1	RP41_METJA	O58445	methanococc	501	5	1.8	69	1	YVCA_VACCC	P21119	vacinia vi
429	6	2.2	1363	1	ILPR_BRALA	O02466	brachyvirio	502	5	1.8	71	1	KEB1_ECOLI	O52279	escherichia
430	6	2.2	1371	1	VCAP_HSVSA	Q00999	herpesviru	503	5	1.8	71	1	V539_TREPA	P06461	human papil
431	6	2.2	1372	1	INSR_MOUSE	P15208	mus musculus	504	5	1.8	72	1	Y539_TREPA	O83550	treponema p
432	6	2.2	1380	1	CYAA_LEIDO	Q27675	leishmania	505	5	1.8	73	1	Y539_TREPA	O83550	treponema p
433	6	2.2	1383	1	INSR_RAT	P5127	rattus norv	506	5	1.8	75	1	RL4E_METTH	O26139	methanobact
434	6	2.2	1433	1	SUBF_BACSU	P16397	bacillus su	507	5	1.8	75	1	RS18_WIGBR	O8d2a3	wiggleswort
435	6	2.2	1435	1	SMF_HUMAN	Q12766	homo sapien	508	5	1.8	75	1	TRY3_ECOLI	P05835	escherichia
436	6	2.2	1453	1	MLH3_HUMAN	Q9h4c1	homo sapien	509	5	1.8	75	1	YVAF_VACCC	P20515	vacinia vi
437	6	2.2	1476	1	A2M1_MOUSE	P28665	mus musculus	510	5	1.8	76	1	CRP1_HUMAN	P50238	homo sapien
438	6	2.2	1556	1	GLT5_SYNY3	P55038	synchocyst	511	5	1.8	76	1	CRP1_MOUSE	P04006	mus musculus
439	6	2.2	1660	1	VIT6_OSCBR	Q94637	oscheilus br	512	5	1.8	76	1	EX78_LACPL	O88w66	lactobacill
440	6	2.2	1788	1	POLN_SO0V3	Q04544	southampton	513	5	1.8	76	1	Y008_LISIN	Q928z3	listeria in
441	6	2.2	1815	1	SHK3_RAT	O9j1i4	rattus norv	514	5	1.8	77	1	RS9_TOBAC	P49214	nicotiana t
442	6	2.2	1819	1	GCP6_HUMAN	O96r17	homo sapien	515	5	1.8	77	1	TAT1_AQUAE	O66478	aquifex aeo
443	6	2.2	1918	1	KE04_HUMAN	O9P2c4	homo sapien	516	5	1.8	78	1	Y084_TREPA	O8d3f1	wiggleswort
444	6	2.2	2032	1	CTOG_HUMAN	Q14008	homo sapien	517	5	1.8	78	1	RL26_WIGBR	O8d3f1	wiggleswort
445	6	2.2	2035	1	HMP2_YEREN	P48633	yeastinia en	518	5	1.8	78	1	RLX_METTH	O27647	methanobact
446	6	2.2	2168	1	POLG_PEV9U	O41174	p genome po	519	5	1.8	79	1	RPOH_METAC	O8ctb4	methanobact
447	6	2.2	2175	1	POLG_BOVEV	P12915	b genome po	520	5	1.8	79	1	VF07_VARV	O8m9y7	chaetophae
448	6	2.2	2233	1	RRPL_P13H4	P12577	human parai	521	5	1.8	79	1	AKAT_MOUSE	P33867	variola vir
449	6	2.2	2258	1	FA5_PIG	Q9G1s1	sus scrofa	522	5	1.8	80	1	DBHL_RICPR	O55074	mus musculus
450	6	2.2	2339	1	RRCT_PLAFA	P27635	plasmodium	523	5	1.8	80	1	DBHL_RICPR	O92426	ricketsia
451	6	2.2	2469	1	TEGU_HSVSA	Q01056	herpesviru	524	5	1.8	81	1	ACP_THEMA	O9wzd0	thermotoga
452	6	2.2	2701	1	IP3S_HUMAN	Q14571	homo sapien	525	5	1.8	81	1	VSH_BRSV3	P32554	bovine resp
453	6	2.2	2709	1	IP3S_RAT	P29995	rattus norv	526	5	1.8	81	1	CY2_RHOGE	P22895	escherichia
454	6	2.2	2749	1	IP3R_MOUSE	O9t2u4	bos taurus	527	5	1.8	83	1	PAD4_ECOLI	P00097	rhodocycylus
455	6	2.2	2750	1	IP3R_RAT	P29994	rattus norv	528	5	1.8	85	1	MTRE_METTH	O27225	methanobact
456	6	2.2	2758	1	IP3R_HUMAN	Q14643	homo sapien	529	5	1.8	85	1	VGT3_BEWLS	O05288	mycobacteri
457	6	2.2	2778	1	FAF_DROME	P55824	drosophila	530	5	1.8	86	1	ACP_CYACA	O19921	cyanidium c
458	6	2.2	2869	1	RBPI_PLAVB	Q00798	plasmodium	531	5	1.8	86	1	GNZ2_ONCMY	O42241	oncorhynch
459	6	2.2	3562	1	PGCV_CHICK	O90953	gallus galli	532	5	1.8	86	1	GVPA_BACME	O06680	bacillus me
460	6	2.2	3829	1	SACS_HUMAN	O9n214	homo sapien	533	5	1.8	86	1	RS16_BORBU	O51638	borrelia bu
461	6	2.2	3830	1	SACS_MOUSE	Q9j1e8	mus musculus	534	5	1.8	86	1	Y16_BP77	P03793	bacterioph
462	6	2.2	4660	1	LRP2_RAT	P98158	rattus norv	535	5	1.8	86	1	YB64_PASMU	O9v9q9	drosophila
463	6	2.2	5065	1	LRP2_HUMAN	P98107	homo sapien	536	5	1.8	87	1	ATPK_MOUSE	O9c1q3	pasteurella
464	6	2.2	5938	1	MAC4_HUMAN	Q96d42	homo sapien	537	5	1.8	87	1	HBH_DASVI	P56135	mus musculus
465	6	2.2	6669	1	NEBU_HUMAN	P03929	homo sapien	538	5	1.8	87	1	AP21_MOUSE	P07420	dasyurus vi
466	5	1.8	10	1	PAR2_PENMO	P83317	penaeus mon	539	5	1.8	88	1	FR1B_THBAC	O9dcb4	mus musculus
467	5	1.8	21	1	NDK_CANAL	Q9ur66	candida alb	540	5	1.8	88	1	FR16_OENHO	O9hkn1	thermoplaem
468	5	1.8	33	1	ALOX_PICPA	P04842	pichia past	541	5	1.8	88	1	RS15_MYCHY	O9mck0	oenothera h
469	5	1.8	36	1	FARE_METOG	P80951	methanogeni	542	5	1.8	88	1	SAPB_SARPE	P46189	mycoplasma
470	5	1.8	39	1	PD43_PAPHA	P81246	papio hamad	543	5	1.8	88	1	SAPB_SARPE	P31529	sarcophaga
471	5	1.8	39	1	PD43_PAPHA	P81246	papio hamad	544	5	1.8	88	1	SAPB_SARPE	P31529	sarcophaga



545	5	1.8	88	1	Y007_STRPN	Q97tc9 streptococc	618	5	1.8	108	1	ONCO_CAVPO	O35508 cavia porce
546	5	1.8	88	1	Y874_CAUCR	P58091 cauliobacter	619	5	1.8	108	1	ONCO_HUMAN	P32930 homo sapien
547	5	1.8	88	1	Y0C6_BACSU	P7481 cauliobacter su	620	5	1.8	108	1	PRVA_BRARE	P31890 brachydanio
548	5	1.8	89	1	AP21_HUMAN	O9ub10 homo sapien	621	5	1.8	108	1	PRVA_CYPCA	P09227 cyprinus ca
549	5	1.8	89	1	GATC_THETH	O91cx4 thermus the	622	5	1.8	108	1	PRVA_ESOLU	P02628 esox lucius
550	5	1.8	89	1	PSK1_ORISA	O9st21 oryza sativ	623	5	1.8	108	1	PRVA_AMPME	P02616 amphiuma me
551	5	1.8	89	1	TRP6_BOVIN	O6my40 bos taurus	624	5	1.8	108	1	PRVA_GPACE	P02614 cyprinus ca
552	5	1.8	90	1	DBHL_RICRI	O6my40 bos taurus	625	5	1.8	108	1	PRVA_GPACE	P02614 cyprinus ca
553	5	1.8	90	1	IM10_NEUCR	O9c0n3 neurospora	626	5	1.8	108	1	PRVA_MERBT	P02631 merluccius
554	5	1.8	91	1	A9_ABATH	O00762 arabidopsis	627	5	1.8	108	1	PRVA_MERBT	P02631 merluccius
555	5	1.8	92	1	KEPK_CANAL	P46614 candida alb	628	5	1.8	108	1	Y1IK_PASVK	O00336 panicum str
556	5	1.8	92	1	PAK2_ANASP	O8y1x8 anabaena sp	629	5	1.8	109	1	PRVA_FELCA	P80079 felis silve
557	5	1.8	92	1	RS19_BUCAI	P57557 buchnera ap	630	5	1.8	109	1	PRVA_GERSP	P80080 gerbillus s
558	5	1.8	92	1	RS19_BUCAI	O8k954 buchnera ap	631	5	1.8	109	1	PRVA_HUMAN	P20472 homo sapien
559	5	1.8	92	1	RS19_NEIMA	O9j9r6 neisseria m	632	5	1.8	109	1	PRVA_MACFU	P80050 mackia fusc
560	5	1.8	92	1	Y367_BORBU	O51340 borrelia bu	633	5	1.8	109	1	PRVA_MOUSE	P32898 mus musculu
561	5	1.8	93	1	FANA_ECOLI	P07104 escherichia	634	5	1.8	109	1	PRVA_RAT	P02625 rattus norv
562	5	1.8	93	1	VE7_HPV1A	P06465 human papil	635	5	1.8	109	1	PRVA_BOACO	P02615 boa constri
563	5	1.8	94	1	RLJ1_PYRAB	O9uy17 pyrococcus	636	5	1.8	109	1	PRVA_OPSTA	P05941 opsanus tau
564	5	1.8	95	1	BCP4_RABIT	O02247 oryctolagus	637	5	1.8	109	1	PRVA_CHICK	P43305 gallus gall
565	5	1.8	95	1	RLJ1_PYRFU	O8u367 pyrococcus	638	5	1.8	109	1	Y819_METUA	O58229 methanococ
566	5	1.8	95	1	RLJ1_PYRHO	P58189 pyrococcus	639	5	1.8	110	1	PRVA_RANCA	P18087 rana catesb
567	5	1.8	95	1	SL10_PIG	P04163 sus scrofa	640	5	1.8	110	1	YGR8_YEAST	P45820 saccharomyc
568	5	1.8	96	1	PAP5_CARPA	P05993 carica papa	641	5	1.8	111	1	YHO7_YEAST	P38804 saccharomyc
569	5	1.8	96	1	PRP5_HUMAN	P04281 homo sapien	642	5	1.8	112	1	CHUR_CHICK	O9df23 gallus gall
570	5	1.8	96	1	SL10_HUMAN	P08206 homo sapien	643	5	1.8	112	1	GLNB_PORPU	P51254 porphyra pu
571	5	1.8	96	1	SL10_XENTLA	P27004 xenopus lae	644	5	1.8	112	1	PT07_STRYP	P28199 strydoma plic
572	5	1.8	96	1	VE7_HP24	P50782 human papil	645	5	1.8	112	1	RSR1_VIRCH	O34419 vibrio chol
573	5	1.8	96	1	VPR_HVISC	P05951 human immun	646	5	1.8	112	1	YAY4_METUA	O58474 methanococ
574	5	1.8	97	1	FIXX_RHISN	O33207 rhizobium s	647	5	1.8	113	1	PRVA_GADCA	P02622 gadus galla
575	5	1.8	97	1	GATC_SULTO	O971u4 sulfolobus	648	5	1.8	113	1	SMPA_ECOLI	O60389 escherichia
576	5	1.8	97	1	RL21_METMA	O8pu21 methanosarc	649	5	1.8	113	1	Y076_METUA	O26383 methanococ
577	5	1.8	97	1	RS6_LISIN	O92f66 listeria in	650	5	1.8	114	1	FLIA_PYRAB	O8xy90 pyrobactul
578	5	1.8	97	1	RS6_LISMO	O8yarc9 listeria mo	651	5	1.8	115	1	GLN2_METTH	O26760 methanobact
579	5	1.8	98	1	HUPF_BRAJA	O45252 bradyrhizob	652	5	1.8	115	1	SMST_RANRI	P87384 rana ridibu
580	5	1.8	98	1	NULM_XENTLA	P03904 xenopus lae	653	5	1.8	115	1	TKN1_RABIT	P41540 schizosacch
581	5	1.8	98	1	VE7_PCPE1	O02272 pygmy chimp	654	5	1.8	115	1	YAPD_SCHPO	P09789 schizosacch
582	5	1.8	98	1	VG04_BPMLS	O05251 mycobacteri	655	5	1.8	116	1	SMS_CHICK	P33094 gallus gall
583	5	1.8	99	1	THFA_PASHA	P95516 pasteurella	656	5	1.8	116	1	Y1B_SOCMV	P15632 soybean chl
584	5	1.8	99	1	THFA_RHOCA	P30787 rhodobacter	657	5	1.8	116	1	YCFE_HAEIN	P44936 haemophilus
585	5	1.8	99	1	PD11_BOVIN	P56982 bos taurus	658	5	1.8	117	1	CHH3_PENUP	P04676 penaeus jap
586	5	1.8	99	1	PD11_HUMAN	O9y661 homo sapien	659	5	1.8	117	1	V03_BPT7	P28091 bacterioph
587	5	1.8	99	1	PD11_MOUSE	P56983 mus musculu	660	5	1.8	117	1	MNT6_EVATR	P28091 evasterias
588	5	1.8	99	1	PD11_PIG	O57606 methanococ	661	5	1.8	117	1	MNT6_STRPU	P28097 strongyloce
589	5	1.8	99	1	Y141_METUA	P44073 haemophilus	662	5	1.8	117	1	YCC1_YEAST	P25561 saccharomyc
590	5	1.8	100	1	Y908_HAEIN	O9t1c4 cyanidium c	663	5	1.8	118	1	DNB1_POVBO	P24850 bovine poly
591	5	1.8	101	1	RK23_CYACA	P08814 bos taurus	664	5	1.8	118	1	Y023_METUA	O60333 methanococ
592	5	1.8	101	1	THYP_BOVIN	P20962 homo sapien	665	5	1.8	118	1	YAI9_SCHPO	O09896 schizosacch
593	5	1.8	101	1	THYP_HUMAN	P04550 ratius norv	666	5	1.8	118	1	YDIL_ECOLI	P76196 escherichia
594	5	1.8	101	1	THYP_RAT	P27553 human papil	667	5	1.8	119	1	Y14_YERPE	O82b55 yerquina pe
595	5	1.8	101	1	VEA_HPVA1	P37205 thunnus obe	668	5	1.8	120	1	R22A_YEAST	P05749 saccharomyc
596	5	1.8	102	1	GTHI_THUOB	P29627 mesocricetu	669	5	1.8	121	1	CHAV_LYMDI	P50603 lymphocita d
597	5	1.8	102	1	HBA2_MESAU	O62100 mus musculu	670	5	1.8	121	1	RA14_CANAL	P53709 candida alb
598	5	1.8	102	1	HSP3_MOUSE	P19953 halorcula	671	5	1.8	122	1	FLIT_SALTY	P26611 salmonella
599	5	1.8	102	1	RS24_HALMA	P08795 bacterioph	672	5	1.8	122	1	YHIT_AZOB	P26740 azospirillu
600	5	1.8	102	1	VGES_BPDI0	O65944 canine aden	673	5	1.8	123	1	MP13_PRAU	P18152 francisella
601	5	1.8	103	1	HEX9_ADECC	O27943 archaeoglob	674	5	1.8	123	1	PYP_RHOCA	O69138 rhodobacter
602	5	1.8	103	1	YN41_ARCFU	O64256 rattus norv	675	5	1.8	124	1	SY02_RABIT	P28292 oryctolagus
603	5	1.8	104	1	HSP3_RAT	P29447 dictyosteli	676	5	1.8	125	1	SZ09_HUMAN	O07335 homo sapien
604	5	1.8	104	1	THI3_DICDI	O57676 methanococ	677	5	1.8	125	1	VG19_BPMLS	O05225 mycobacteri
605	5	1.8	104	1	Y223_METUA	O83278 creponema p	678	5	1.8	125	1	Y3C2_STRCO	O53968 streptomyce
606	5	1.8	105	1	DBH_TBREA	O95178 homo sapien	679	5	1.8	125	1	YCK1_GUTH	O78420 guilardia
607	5	1.8	105	1	NIGM_HUMAN	P38513 thermotoga	680	5	1.8	125	1	YCK8_OENHO	O9m4m7 oenothera h
608	5	1.8	105	1	RL24_THEMA	O92cc0 rickettsia	681	5	1.8	126	1	C1X_HUMAN	O15332 homo sapien
609	5	1.8	105	1	Y839_RICPR	O92g24 rickettsia	682	5	1.8	126	1	CRGB_VIRCH	O9kvs9 vibrio chol
610	5	1.8	106	1	YD01_RICCN	O60397 homo sapien	683	5	1.8	126	1	RECK_HAEIN	P43706 haemophilus
611	5	1.8	106	1	C07S_HUMAN	P43566 saccharomyc	684	5	1.8	127	1	YCBK_HAEIN	P44284 haemophilus
612	5	1.8	107	1	YFD2_YEAST	O89a15 buchnera ap	685	5	1.8	129	1	HVZF_HUMAN	P01924 homo sapien
613	5	1.8	107	1	PER_BUCEP	O91483 salmo salar	686	5	1.8	129	1	RR8_NEPOL	O9t124 neptroselimi
614	5	1.8	107	1	PRVB_SALSA	P02619 esox lucius	687	5	1.8	129	1	TKN1_HUMAN	P02036 homo sapien
615	5	1.8	107	1	THIO_CLOLI	O23747 clostridium	688	5	1.8	130	1	INIT_HUMAN	P09912 homo sapien
616	5	1.8	108	1	NIGM_BOVIN	O02374 bos taurus	689	5	1.8	130	1	INIT_PANTR	O28808 pan troglod
617	5	1.8	108	1	NIGM_BOVIN	O02374 bos taurus	690	5	1.8	130	1	INIT_PANTR	O28808 pan troglod

691	5	1.8	130	1	TKN1_MESAU	060541 mesocricetu	764	5	1.8	145	1	HBB_TRAST	P04245 tragephaphus
692	5	1.8	130	1	TKN1_MOUSE	P41539 mus musculu	765	5	1.8	145	1	LIRP_LOCM1	P15131 locusta mig
693	5	1.8	130	1	TKN1_RAT	P06767 rattus norv	766	5	1.8	145	1	RL2A_TRYBB	O15683 trypanosoma
694	5	1.8	130	1	YAJI_RHISN	P55509 rhizobium s	767	5	1.8	145	1	URI_CARAU	O9PC94 carassius a
695	5	1.8	130	1	YEM5_YEAST	P40082 saccharomyc	768	5	1.8	145	1	URI_CYPCA	P01146 cyprinus ca
696	5	1.8	130	1	YGB5_YEAST	P33199 saccharomyc	769	5	1.8	145	1	Y004_TREPA	O83050 treponema p
697	5	1.8	130	1	YNS5_CABEI	Q22288 caenorhabdi	770	5	1.8	146	1	CHA3_LYMDI	P433513 lymantria d
698	5	1.8	131	1	CHRI_HUMAN	Q09100 homo sapien	771	5	1.8	146	1	DTD_BACSU	O32042 bacillus su
699	5	1.8	131	1	H2A1_YEAST	P04911 saccharomyc	772	5	1.8	146	1	DUT_ZYMMO	O9J3X5 zymomonas m
700	5	1.8	131	1	H2A2_YEAST	P04912 saccharomyc	773	5	1.8	146	1	GGD2_HUMAN	O9H646 homo sapien
701	5	1.8	131	1	LGUL_SYNY3	Q55595 synechocyst	774	5	1.8	146	1	GPR7_MOUSE	P49581 mus musculu
702	5	1.8	131	1	Y044_NPVAC	P41494 autographa	775	5	1.8	146	1	HBB0_MOUSE	P04443 mus musculu
703	5	1.8	131	1	Y624_MYCTU	P69914 mycobacteri	776	5	1.8	146	1	HBB1_SPHPU	P10060 sphenodon p
704	5	1.8	131	1	YF69_SCHPO	O13929 schizosacch	777	5	1.8	146	1	HBB2_VAREX	P18993 varanus exa
705	5	1.8	132	1	PRT_ANTGR	P17502 anthronomus	778	5	1.8	146	1	HBB3_NAJNA	P22743 najia najia f
706	5	1.8	132	1	SPH2_AERPE	O39737 aeropyrum p	779	5	1.8	146	1	HBB4_MESAU	P29652 mesocricetu
707	5	1.8	132	1	Y513_ARCFU	O29737 archaeoglob	780	5	1.8	146	1	HBB5_MOUSE	P04444 mus musculu
708	5	1.8	133	1	CHAB_LYMDI	P50602 lymantria d	781	5	1.8	146	1	HBB6_AILFU	P18982 ailurus ful
709	5	1.8	133	1	RS19_ARCFU	O28358 archaeoglob	782	5	1.8	146	1	HBB7_AILME	P18983 ailuropoda
710	5	1.8	133	1	RSBT_BACSU	P42411 bacillus su	783	5	1.8	146	1	HBB8_ALDEL	P83133 alidabrachei
711	5	1.8	133	1	Y497_METUA	O57920 methanococc	784	5	1.8	146	1	HBB9_BRATP	P14526 bradyus tr
712	5	1.8	134	1	CLX2_NARJA	O42105 narke japon	785	5	1.8	146	1	HBB10_CANPA	P20256 canis fami1
713	5	1.8	134	1	FLGC_YEREN	Q56894 yerstina en	786	5	1.8	146	1	HBB11_CERAB	P02028 cercopithec
714	5	1.8	134	1	LEGI_SHEEP	P81164 ovib aries	787	5	1.8	146	1	HBB12_CERTO	P02031 cercocebus
715	5	1.8	134	1	VG24_BPMD2	O64218 mycobacteri	788	5	1.8	146	1	HBB13_CHRPI	P13274 chrysomys p
716	5	1.8	134	1	YAAI_ECOLI	P28636 escherichia	789	5	1.8	146	1	HBB14_COLPA	P02033 colobus bad
717	5	1.8	134	1	YN06_PARDE	P29912 paracoccusu	790	5	1.8	146	1	HBB15_COLPO	P19885 colobus pol
718	5	1.8	135	1	C550_MICAE	P19129 microcystis	791	5	1.8	146	1	HBB16_GROCR	P18986 crocuta cro
719	5	1.8	135	1	FLGN_BUCAI	P57417 buchiera ap	792	5	1.8	146	1	HBB17_GEOCR	O98905 geocheilone
720	5	1.8	135	1	HNS_BUCAI	P57360 buchiera ap	793	5	1.8	146	1	HBB18_GEONI	P83123 geocheilone
721	5	1.8	135	1	PGS1_RABIT	O46371 oncorhynch	794	5	1.8	146	1	HBB19_GORGO	P02024 gorilla gor
722	5	1.8	135	1	PPOL_ONCMA	O08824 oncorhynch	795	5	1.8	146	1	HBB20_LEMCA	P02024 lemur catla
723	5	1.8	135	1	Y707_ARCFU	O28014 archaeoglob	796	5	1.8	146	1	HBB21_LEMVA	P21657 lemur varie
724	5	1.8	136	1	RS11_LEBIN	O9X141 leptospira	797	5	1.8	146	1	HBB22_LEPME	P15166 leptonchoc
725	5	1.8	136	1	RS11_MYCPU	Q98Q77 mycoplasma	798	5	1.8	146	1	HBB23_LUTLU	P10893 lutra lutra
726	5	1.8	136	1	YG10_PYPAB	Q9UY98 pyrococcus	799	5	1.8	146	1	HBB24_MACFU	P02027 macaca fusc
727	5	1.8	136	1	YGHG_ECOLI	Q46835 escherichia	800	5	1.8	146	1	HBB25_MACGI	P02106 macropus gi
728	5	1.8	136	1	Y101_CVBM	P22053 bovine coro	801	5	1.8	146	1	HBB26_MACMU	P02026 macaca mula
729	5	1.8	137	1	P16B_CUCMA	Q9Z146 cucurbita m	802	5	1.8	146	1	HBB27_MACCRU	P02107 macropus ru
730	5	1.8	137	1	PSBW_SPIOL	O41387 spinacia ol	803	5	1.8	146	1	HBB28_MANSP	P08853 marmotilla
731	5	1.8	137	1	RS19_NEPOL	Q9LXK4 nephrolepis	804	5	1.8	146	1	HBB29_MARMA	P08853 marmota mar
732	5	1.8	138	1	PSAD_FREDI	P23808 tremyella d	805	5	1.8	146	1	HBB30_MELCA	P15449 mellivora c
733	5	1.8	138	1	RL16_MYCBO	O06049 mycobacteri	806	5	1.8	146	1	HBB31_MEIME	P02055 melospiza m
734	5	1.8	138	1	RL16_MYCLE	O32968 mycobacteri	807	5	1.8	146	1	HBB32_MESAU	P02094 mesocricetu
735	5	1.8	138	1	RL16_MYCTU	P95056 mycobacteri	808	5	1.8	146	1	HBB33_MESBR	P18707 mesocricetu
736	5	1.8	139	1	ATPD_NEICR	P56525 neurospora	809	5	1.8	146	1	HBB34_MUSLU	P23602 muscicetu
737	5	1.8	140	1	CHAZ_LYMDI	P43512 lymantria d	810	5	1.8	146	1	HBB35_MUSPF	P19017 muscicetu
738	5	1.8	142	1	CHAF_LYMDI	P43516 lymantria d	811	5	1.8	146	1	HBB36_ODORO	P10779 odonemus to
739	5	1.8	142	1	FOB1_STABP	P59231 staphylococ	812	5	1.8	146	1	HBB37_ONDZI	P02093 ondatra zib
740	5	1.8	142	1	MRAZ_CLOAB	Q97860 clostridium	813	5	1.8	146	1	HBB38_PAGIA	P19646 paguma lary
741	5	1.8	142	1	PTA1_XENLA	P48530 xenopus lae	814	5	1.8	146	1	HBB39_PARCY	P02030 papio cynoc
742	5	1.8	142	1	PTA2_XENLA	P48531 xenopus lae	815	5	1.8	146	1	HBB40_POHOT	P02030 papio cynoc
743	5	1.8	142	1	RL13_PYPAB	Q9V196 pyrococcus	816	5	1.8	146	1	HBB41_POTTR	P02108 potoronus tr
744	5	1.8	142	1	RL13_PYPAB	O59133 pyrococcus	817	5	1.8	146	1	HBB42_PROCR	P23020 proteoles cr
745	5	1.8	142	1	Y441_MYCTU	O50813 mycobacteri	818	5	1.8	146	1	HBB43_PTEBR	P10886 pteronura b
746	5	1.8	142	1	CHAS_LYMDI	P43514 lymantria d	819	5	1.8	146	1	HBB44_SPAEH	P02090 spalax leuc
747	5	1.8	143	1	DUT_GLODI	O30931 clostridium	820	5	1.8	146	1	HBB45_SPECI	P09421 spermophilu
748	5	1.8	143	1	FRDA_ARATH	Q9Z107 arabidopsis	821	5	1.8	146	1	HBB46_TADBR	P07410 tadaria br
749	5	1.8	143	1	PSGD_HUMAN	Q9UQ72 homo sapien	822	5	1.8	146	1	HBB47_THEGE	P02022 theropithec
750	5	1.8	143	1	PK13_GUTTH	O46915 guillardi	823	5	1.8	146	1	HBB48_URMSA	P07422 ursus marit
751	5	1.8	144	1	RL15_CHLPN	Q9Z764 chlamydia p	824	5	1.8	146	1	HBB49_VULVU	P21201 vulpes vulp
752	5	1.8	144	1	RL15_SUISO	O9UX85 sulfobobus	825	5	1.8	146	1	HBB50_CHEME	O28338 cheilrogaleu
753	5	1.8	144	1	YDB1_PYPFU	Q9U148 pyrococcus	826	5	1.8	146	1	HBB51_EULFU	P08823 eulemur ful
754	5	1.8	144	1	DPE3_MOUSE	Q9JXK7 mus musculu	827	5	1.8	146	1	HBB52_GALCR	P19795 galago cras
755	5	1.8	145	1	DTD_PSEAE	Q9HNA4 pseudomomas	828	5	1.8	146	1	HBB53_HUMAN	P02100 homo sapien
756	5	1.8	145	1	DUT_HELPU	Q9ZKX3 helicobacte	829	5	1.8	146	1	HBB54_HYLSY	O95190 hylobates s
757	5	1.8	145	1	HBB5A_BOSJA	P04336 bos javanic	830	5	1.8	146	1	HBB55_MACMU	O28507 macaca mula
758	5	1.8	145	1	HBB5A_CAPHI	P02077 capra hircu	831	5	1.8	146	1	HBB56_MICMU	O28496 microtus
759	5	1.8	145	1	HBB5B_BISBO	P09422 bison bonas	832	5	1.8	146	1	HBB57_MOUSE	P02104 mus musculu
760	5	1.8	145	1	HBB5B_BOSGP	P02071 bos gaurus	833	5	1.8	146	1	HBB58_PANPA	O28775 pan paniscu
761	5	1.8	145	1	HBB5B_BOSMU	P02072 bos mutus g	834	5	1.8	146	1	HBB59_PONPY	P18994 pongo pygma
762	5	1.8	145	1	HBB5B_BOVIN	P02070 bos taurus g	835	5	1.8	146	1	HBB60_PROVE	O95238 propithecus
763	5	1.8	145	1	HBB5B_BOVIN	P02070 bos taurus g	836	5	1.8	146	1	HBB61_PROVE	O95238 propithecus

837	5	1.8	146	1	HBB_RABIT	P02103	oryctolagus	910	5	1.8	157	1	BSAA_BACHID	Q929n7	bacillus ha
838	5	1.8	146	1	HBG_CHEMIE	P08224	cheilrogaleu	911	5	1.8	157	1	MNBB_MAIZE	P27347	zea mays (m
839	5	1.8	146	1	HBG_EULFJ	P08225	elenmurr ful	912	5	1.8	157	1	VE6_HPV23	P50776	human papil
840	5	1.8	146	1	HBG_HUMAN	P02096	homo sapien	913	5	1.8	157	1	Y935_AQUAE	O67076	aquifex aeo
841	5	1.8	146	1	HBG_RABIT	P02099	oryctolagus	914	5	1.8	157	1	YBUX_CLOAB	O05526	clostridium
842	5	1.8	146	1	MIOC_ECO57	P58208	eschericchia	915	5	1.8	157	1	YF53_METUA	O05894	methanococ
843	5	1.8	146	1	RL32_METUA	P54010	methanococ	916	5	1.8	157	1	YMFO_ECOLI	P75980	eschericchia
844	5	1.8	147	1	GLSH_AERPE	O9ydg2	aeropyrum p	917	5	1.8	158	1	BFR_ECOLI	P11056	eschericchia
845	5	1.8	147	1	SSB_AQUAE	O66475	aquifex aeo	918	5	1.8	158	1	BFG2_RAT	O68926	salmonella
846	5	1.8	147	1	TPC_TODPA	O9b1g0	todarodes p	919	5	1.8	158	1	COAD_PYRHO	P27049	rattus norv
847	5	1.8	147	1	Y521_ARCFU	O29729	archaeoglob	920	5	1.8	158	1	FCY1_YEAST	O58358	pyrococcus
848	5	1.8	147	1	YLVA_BACSU	O45478	baclillus su	921	5	1.8	158	1	GSHD_SCHPO	Q12178	saccharomyc
849	5	1.8	147	1	YZ20_METUA	O60280	methanococ	922	5	1.8	158	1	RL15_AERPE	O59688	schizosacch
850	5	1.8	148	1	NEU2_CATCO	P80722	isoparorchl	923	5	1.8	158	1	RL15_AERPE	O9y198	aeropyrum p
851	5	1.8	148	1	GLB_ISOBY	P15211	calotomus	924	5	1.8	158	1	RS19_PYRAE	O8zu14	pyrobaculum
852	5	1.8	149	1	CTH3_HUMAN	O961m9	homo sapien	925	5	1.8	158	1	Y4A8_METUA	O57912	methanococ
853	5	1.8	149	1	HMGL_VICFA	P40620	vicia faba	926	5	1.8	159	1	ATPX_ANTSP	O02852	methanococ
854	5	1.8	149	1	NUSB_CACCR	O9a8j3	caulobacter	927	5	1.8	159	1	BFR_MYCLE	P43315	mycobacteri
855	5	1.8	149	1	P16A_CUCMA	O9z147	cucurbita m	928	5	1.8	159	1	BFR_MYCTU	O08465	mycobacteri
856	5	1.8	149	1	Y1L1_STNAU	P13369	staphylococ	929	5	1.8	159	1	BV1E_SERMA	O68935	serratia ma
857	5	1.8	149	1	Y547_CABEL	O09373	caenorhabdi	930	5	1.8	159	1	BV1E_BETVE	P43178	betula verr
858	5	1.8	150	1	ARG2_CLOPE	P58685	schizosacch	931	5	1.8	159	1	BV1F_BETVE	P43183	betula verr
859	5	1.8	150	1	CALM_SCHPO	P05933	schizosacch	932	5	1.8	159	1	BV1F_BETVE	P43178	betula verr
860	5	1.8	150	1	DEPA_ASTAM	P20630	asterias am	933	5	1.8	159	1	RIB4_SCHPO	O9ubd1	schizosacch
861	5	1.8	150	1	DTD_LISIN	O92b11	listeria in	934	5	1.8	159	1	Y215_METUA	O57668	methanococ
862	5	1.8	150	1	DUT_AQUAE	O66592	aquifex aeo	935	5	1.8	160	1	C550_SYNY3	O55013	synchocyst
863	5	1.8	150	1	DUT_NEIMA	O91un1	neisseria m	936	5	1.8	160	1	CNTL_HUMAN	O8tbel	homo sapien
864	5	1.8	150	1	DUT_NEIMB	O91zu7	neisseria m	937	5	1.8	160	1	CNTL_MOUSE	O35089	mus musculu
865	5	1.8	150	1	RRF2_DESVH	P33395	desulfovibri	938	5	1.8	160	1	COAD_PYRFU	O8u1u0	pyrococcus
866	5	1.8	150	1	TCPO_VIBCH	P29450	vibrio chol	939	5	1.8	160	1	IAGB_SALTY	P43018	salmonella
867	5	1.8	150	1	TPC1_HOMAM	P29289	homarus ame	940	5	1.8	160	1	IAGB_SALTY	P43017	salmonella
868	5	1.8	150	1	TPC1_PONLE	P06707	pontastacus	941	5	1.8	160	1	MAX_CHICK	P52162	gallus gall
869	5	1.8	150	1	TPIS_MYCPI	P47721	mycoplasma	942	5	1.8	160	1	MAX_HUMAN	P25912	gallus gall
870	5	1.8	150	1	Y109_CHLMU	O9p1j0	chlamydia m	943	5	1.8	160	1	MAX_MOUSE	P26574	mus musculu
871	5	1.8	150	1	Y736_CHLTR	O84741	chlamydia t	944	5	1.8	160	1	MAX_RAT	P52164	rattus norv
872	5	1.8	150	1	Y877_CHLPN	O9z729	chlamydia p	945	5	1.8	161	1	HMGL_WHEAT	P40621	trititum ae
873	5	1.8	150	1	YDPC_SCHPO	O14015	schizosacch	946	5	1.8	161	1	TPCC_CHICK	P09860	gallus gall
874	5	1.8	150	1	YPSL_SYNEN	P25908	synchococ	947	5	1.8	161	1	TPCC_COTUA	P05396	cornutrix co
875	5	1.8	151	1	CAV3_HUMAN	P56539	homo sapien	948	5	1.8	161	1	TPCC_HUMAN	P02590	homo sapien
876	5	1.8	151	1	CAV3_MOUSE	P51637	mus musculu	949	5	1.8	161	1	TPCC_MOUSE	P19123	mus musculu
877	5	1.8	151	1	CAV3_RAT	P51638	rattus norv	950	5	1.8	161	1	TPCC_RABIT	P02591	oryctolagus
878	5	1.8	151	1	MUP8_MOUSE	P04938	mus musculu	951	5	1.8	161	1	TPM_SCHPO	O02088	schizosacch
879	5	1.8	151	1	NRD1_MYCGA	O9xc21	mycoplasma	952	5	1.8	161	1	YU12_CLOPE	O93md5	clostridium
880	5	1.8	151	1	SMD3_DROME	O44373	methanobact	953	5	1.8	162	1	FLGA_AGRFS	O44339	agrobacteri
881	5	1.8	151	1	YZ73_METHU	O26373	methanobact	954	5	1.8	162	1	PHZB_PSEPL	O51788	pseudomonas
882	5	1.8	152	1	RL30_ARCFU	O28375	archaeoglob	955	5	1.8	162	1	PTSN_KLEBN	P17162	klebsiella
883	5	1.8	152	1	RS19_METUA	P54018	methanococ	956	5	1.8	162	1	RECK_XANAC	O91cz3	xanthomonas
884	5	1.8	152	1	Y4UB_RH1SN	P55502	rhizobium s	957	5	1.8	162	1	RECK_XANOR	O9a1c3	xanthomonas
885	5	1.8	152	1	Y533_CHLPN	O9z819	chlamydia p	958	5	1.8	162	1	YBPM_BACSU	O31453	bacillus su
886	5	1.8	152	1	Y686_CHLMU	O9p1j0	chlamydia m	959	5	1.8	163	1	IPYR_STRCO	O9x819	streptococ
887	5	1.8	153	1	PYR1_VIBPA	O87f57	vibrio para	960	5	1.8	163	1	LEUD_PYRHO	O59393	pyrococcus
888	5	1.8	153	1	Y156_UREPA	O9p4y8	ureaplasma	961	5	1.8	163	1	MAX_XENIA	O07016	xenopus lae
889	5	1.8	153	1	YH17_YEAST	P31898	saccharomyc	962	5	1.8	163	1	MOAE_BRUME	O8y9a6	bruceella m
890	5	1.8	153	1	YZ30_METUA	O60288	methanococ	963	5	1.8	163	1	PHAC_CYACA	O9ltu4	cyathidium c
891	5	1.8	154	1	BFRA_NEIMA	P56998	neisseria m	964	5	1.8	163	1	PTSN_ECOLI	P31222	eschericchia
892	5	1.8	154	1	BFRA_NEIMB	P72080	neisseria m	965	5	1.8	163	1	Y057_NPVOP	O10315	orygia pseu
893	5	1.8	154	1	KMLS_METGA	P56276	meleagris g	966	5	1.8	163	1	YAG5_SCHPO	O09669	schizosacch
894	5	1.8	154	1	NEU1_CATCO	P15210	calotomus	967	5	1.8	164	1	RPOB_BACHD	O9Kcd6	baclillus ha
895	5	1.8	154	1	PYR1_HALNI	O9h1n3	halobacteri	968	5	1.8	165	1	GCSH_DROME	O9u616	drosophila
896	5	1.8	154	1	RL30_METUA	P54016	methanococ	969	5	1.8	165	1	MAX_BRARE	P52161	brachydanto
897	5	1.8	154	1	RL11_HUMAN	O9y3c5	homo sapien	970	5	1.8	165	1	PDAL_METMA	P58889	methanosarc
898	5	1.8	154	1	RN11_MOUSE	O9qyK7	mus musculu	971	5	1.8	165	1	PSAF_SYNY3	P29256	synchocyst
899	5	1.8	154	1	Y406_CHLTR	O84411	chlamydia t	972	5	1.8	165	1	YC89_METUA	O58685	methanococ
900	5	1.8	154	1	YX2_CVHSA	P59633	human coron	973	5	1.8	166	1	PMFA_CANBO	P14592	candida boi
901	5	1.8	155	1	C550_PARDE	P00096	paracoccus	974	5	1.8	166	1	PMFB_CANBO	P15193	nocardia boi
902	5	1.8	155	1	NUSB_VIBFI	O8g9g3	vibrio fisc	975	5	1.8	166	1	VSNI_NOCAR	P51048	nocardia ae
903	5	1.8	155	1	NUSB_VIBVU	O8df28	aquifex aeo	976	5	1.8	166	1	Y346_MYCGE	P47588	mycoplasma
904	5	1.8	155	1	YB88_AQUAE	O67240	aquifex aeo	977	5	1.8	166	1	Y346_MYCPN	P75557	mycoplasma
905	5	1.8	156	1	ATPX_ODOST	O00823	odontella s	978	5	1.8	166	1	YC93_CLOAB	O97j17	clostridium
906	5	1.8	156	1	BFR_SYNY3	P24602	synchocyst	979	5	1.8	167	1	MB12_MYCLE	O49717	mycobacteri
907	5	1.8	156	1	COAD_PYRAB	O9uyt0	pyrococcus	980	5	1.8	167	1	T2DB_YEAST	P11747	saccharomyc
908	5	1.8	156	1	LUXS_STAPB	O8cm10	staphylococ	981	5	1.8	167	1	Y199_MYCGE	P47441	mycoplasma
909	5	1.8	156	1	NUSB_VIBCH	O9kp45	vibrio chol	982	5	1.8	168	1	CDN6_HUMAN	P42773	homo sapien

983	5	1.8	168	1	CDN6_MOUSE	O60772	mus musculus
984	5	1.8	168	1	MB12_MYCTU	O05822	mycobacteri
985	5	1.8	169	1	CHE1_ANTRO	P05664	antrepera p
986	5	1.8	169	1	Y538_NEIMB	O9K0D7	neisseria m
987	5	1.8	170	1	VS11_ROTGA	P18571	rotavirus m
988	5	1.8	170	1	YCS1_YEAST	P18571	rotavirus m
989	5	1.8	170	1	YCS1_YEAST	P18571	rotavirus m
990	5	1.8	171	1	CD18_ARATH	P42951	saccharomy
991	5	1.8	171	1	CD18_ARATH	P42951	saccharomy
992	5	1.8	171	1	CD18_ARATH	P42951	saccharomy
993	5	1.8	171	1	CD18_ARATH	P42951	saccharomy
994	5	1.8	172	1	Y117_NEIMA	O96K39	homo sapien
995	5	1.8	172	1	CATR_NAERG	O93441	naegleria m
996	5	1.8	172	1	CATR_NAERG	O93441	naegleria m
997	5	1.8	173	1	FX12_BANCA	O91W29	mus musculu
998	5	1.8	174	1	FX12_BANCA	O91W29	mus musculu
999	5	1.8	174	1	FX12_BANCA	O91W29	mus musculu
1000	5	1.8	174	1	FX12_BANCA	O91W29	mus musculu

## ALIGNMENTS

ID	MYB_DNA_BINDING	STANDARD;	PRT;	252 AA.
AC	P23592;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Anthocyanin regulatory C1-I protein.			
GN	C1.			
OS	Zea mays (Maize).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	NCCAD clade; Panicoideae; Andropogoneae; Zea.			
OK	PNC1_TaxID=4577;			
RP	[1]			
RL	SEQUENCE FROM N.A.			
RX	MEDLINE=90151603; PubMed=2303027;			
RA	Paz-Ares J., Ghosal D., Saedler H.;			
RT	"Molecular analysis of the C1-I allele from Zea mays: a dominant			
RT	mutant of the regulatory C1 locus.";			
RL	EMBO J. 9:315-321(1990).			
CC	-I- FUNCTION: CONTROLS THE EXPRESSION OF GENES INVOLVED IN ANTHOCYANIN			
CC	BIOSYNTHESIS. REGULATES THE EXPRESSION OF AT LEAST 3 STRUCTURAL			
CC	GENES: CHALCONE SYNTHASE, DIHYDROFLAVONOL REDUCTASE AND FLAVONOL			
CC	O(3) GLUCOSYLTRANSFERASE. C1 ACTS AS A TRANS-ACTING FACTOR.			
CC	-I- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	-I- MISCELLANEOUS: THE C1-I ALLELE IS A DOMINANT NEGATIVE MUTANT WHICH			
CC	INHIBITS PIGMENT FORMATION.			
CC	-I- SIMILARITY: Contains 2 Myb-like domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce">http://www.isb-sib.ch/announce</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; X52201; CAA36456.1; -			
DR	HSSP; P01103; IPOM.			
DR	TRANSFAC; T02946; -			
DR	MaizeDB; 24964; -			
DR	InterPro; IPR001005; MYB_DNA_BINDING.			
DR	Pfam; PF00247; myb_DNA-Binding; 2.			
DR	SMART; SM00719; SANT; 2.			
DR	PROSITE; PS00037; MYB_1; 1.			
DR	PROSITE; PS00334; MYB_2; 1.			
DR	PROSITE; PS00900; MYB_3; 2.			
DR	Nuclear protein; DNA-binding; Repeat; Transcription regulation;			
DR	Activator; Trans-acting factor.			

FT	DNA_BIND	9	61	MYB 1.
FT	DNA_BIND	62	112	MYB 2.
FT	DOMAIN	203	212	POLY-GLY.
SO	SEQUENCE	252 AA;	26419 MM;	AD61FA2B829C5FBD CRC64;

Query Match Similarity	6.2%	Score 17	DB 1	length 252
Best Local Similarity	100.0%	Pred. No.	6.7e-10	
Matches 17	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	83	LLGNRMSLIAGRLPGRT	99
Db	84	LLGNRMSLIAGRLPGRT	100

ID	TT2_ARATH	STANDARD;	PRT;	258 AA.
CC	TT2_ARATH			
AC	OPRTA2:			
DT	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	TRANSFERRIN TESTA 2 protein (Myb-related protein 123) (AtMYB123) (Myb-			
DE	related transcription factor LBM2-like).			
GN	TT2 OR MYB123 OR ATMS35550 OR MOK9.18.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyte; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosids II; Brassicales; Brassicaceae; Arabidopsids.			
OX	NCBI_Taxid=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A. AND MUTANTS TT2-2 AND TT2-4.			
RC	STRAIN=cv. Columbia, cv. Landsberg erecta, and cv. Wassilewskija;			
RX	MEDLINE=21434422; PubMed=1159766;			
RX	Medi N., Jond C., Debeunon I., Caboche M., Lepoint L.;			
RT	"The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts			
RT	as a key determinant for proanthocyanidin accumulation in developing			
RT	seed.";			
RL	Plant Cell 13:2099-2114(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=21481677; PubMed=11597504;			
RA	Stracke R., Werder M., Weisshaar B.;			
RT	"The R2R3-MYB gene family in Arabidopsis thaliana.";			
RL	Curr. Opin. Plant Biol. 4:447-456(2001).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=99087489; PubMed=9872454;			
RA	Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,			
RA	Tabata S.;			
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. VII.			
RT	Sequence features of the regions of 1,013,767 bp covered by sixteen			
RT	physically assigned P1 and YAC clones.";			
RL	DNA Res. 5:297-308(1998).			
CC	-1- FUNCTION: Transcription factor involved in the control of			
CC	flavonoid late metabolism in developing siliques. Plays a key role			
CC	in determining the tissue-specific activation of			
CC	leucoanthocyanidin reductase (BANYSUS).			
CC	-1- SUBCELLULAR LOCATION: Nuclear.			
CC	-1- TISSUE SPECIFICITY: Expressed at a high level in immature siliques			
CC	and at a lower level in flowers. Undetected in young seedlings,			
CC	roots, leaves and inflorescence stems.			
CC	-1- DEVELOPMENTAL STAGE: Highly expressed from the very early stages			
CC	of embryogenesis to the globular stage, decreases rapidly from			
CC	the late heart-shaped stage and did not persist after the			
CC	completion of embryogenesis.			
CC	-1- MISCELLANEOUS: TT2 activity is tightly linked to the presence of			
CC	TT8.			
CC	-1- SIMILARITY: Contains 2 Myb-like domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 CC EMBL; AJ299452; CAC40021.1; -  
 DR EMBL; AF371981; AAK54744.1; -  
 DR EMBL; AB015477; BAB08716.1; -  
 DR HSSP; P06876; 1MBK.  
 DR InterPro; IPR001005; Myb DNA-binding.  
 DR Pfam; PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KM Flavonoid biosynthesis; Transcription; Transcription regulation;  
 KM Nuclear protein; ATP-binding; DNA-binding; Repeat; Polymorphism.  
 FT DNA\_BIND 11 63 MYB 1.  
 FT NP\_BIND 64 114 MYB 2.  
 FT NP\_BIND 47 54 ATP (POTENTIAL).  
 FT VARIANT 174 174 O -> L (IN CV. WS).  
 FT MUTAGEN 66 66 G->R: IN TT2-4; LOSS OF SEED  
 FT PIGMENTATION.  
 FT MUTAGEN 127 258 MISSING: IN TT2-2; REDUCED SEED  
 FT PIGMENTATION.  
 SQ SEQUENCE 258 AA; 29611 MW; 3BA9CB3F723D1C2E CRC64;

Query Match 6.2%; Score 17; DB 1; Length 258;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-10;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 ILGNRWSLIAGRLPGRT 99  
 Db 86 ILGNRWSLIAGRLPGRT 102

RESULT 3  
 MYBC\_MAIZE STANDARD; PRT; 273 AA.  
 AC P10250;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Anthocyanin regulatory Cl protein.  
 GN Cl.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 OC NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88111545; PubMed=3428265;  
 RA Paz-Ares J., Ghosal D., Wienand U., Peterson P.A., Siedler H.;  
 RT "The regulatory cl locus of Zea mays encodes a protein with homology  
 RT to myb proto-oncogene products and with structural similarities to  
 RT transcriptional activators."  
 RL EMBL J. 6:3553-3558(1987).  
 CC -1- FUNCTION: CONTROLS THE EXPRESSION OF GENES INVOLVED IN ANTHOCYANIN  
 CC BIOSYNTHESIS. REGULATES THE EXPRESSION OF AT LEAST 3 STRUCTURAL  
 CC GENES: CHALCONE SYNTHASE, DIHYDROFLAVONOL REDUCTASE AND FLAVONOL  
 CC O(3) GLUCOSYLTRANSFERASE. CI ACTS AS A TRANS-ACTING FACTOR.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 2 Myb-like domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; M37153; AAA33482.1; -  
 DR PIR; S06215; TVZMMB.  
 DR HSSP; P01103; 1POM.  
 DR TRANSFAC; T01592; -.  
 DR MaizeDB; 24964; -.  
 DR InterPro; IPR001005; Myb DNA binding.  
 DR Pfam; PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KM Nuclear protein; DNA-binding; Repeat; Transcription regulation;  
 KM Activator; Trans-acting factor.  
 FT DNA\_BIND 9 61 MYB 1.  
 FT DNA\_BIND 62 112 MYB 2.  
 FT DOMAIN 205 212 POLY-GLY  
 FT DOMAIN 234 273 ASP-GLU-RICH (ACIDIC).  
 SQ SEQUENCE 273 AA; 28750 MW; 163977BBAC6E69D CRC64;

Query Match 6.2%; Score 17; DB 1; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 7.2e-10;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 ILGNRWSLIAGRLPGRT 99  
 Db 84 ILGNRWSLIAGRLPGRT 100

RESULT 4  
 GL1\_ARATH STANDARD; PRT; 228 AA.  
 AC P27900; O95T42;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Trichome differentiation protein G11 (GLABROUS1 protein).  
 DE GL1 OR AT3G27920 OR K16N12.17.  
 GN Arabidopsis thaliana (Mouse-ear cress).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92034971; PubMed=1934056;  
 RA Oppenheimer D.G., Herman P.L., Sivakumaran S., Esch J., Marks M.D.;  
 RT "A myb gene required for leaf trichome differentiation in Arabidopsis  
 RT is expressed in stipules."  
 RL Cell 67:483-493(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=cv. Columbia;  
 RX MEDLINE=99449058; PubMed=10520748;  
 RA Shikazono N., Tanaka A., Yokota Y., Watanabe H., Tano S.;  
 RT "Nucleotide sequence of the GLABROUS1 gene of Arabidopsis thaliana  
 RT ecotype Columbia.";  
 RL DNA Seq. 9:1177-181(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=cv. Columbia, cv. Ba-1, cv. Bla-1, cv. Cond, cv. Sha, and  
 RX cv. Te-0;  
 RX MEDLINE=21396402; PubMed=11504855;  
 RA Hauber M.T., Harr B., Schlotterer C.;  
 RT "Trichome distribution in Arabidopsis thaliana and its close relative  
 RT Arabidopsis lyrata: molecular analysis of the candidate gene  
 RT GLABROUS1.";  
 RL Mol. Biol. Evol. 18:1754-1763(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=cv. Columbia;  
 RX MEDLINE=20563099; PubMed=10907853;  
 RA Kaneke T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.  
RT Sequence features of the 4,251,695 bp regions covered by 90 pl, TAC  
RT DNA clones.":  
RL DNA Res. 7:217-221(2000).  
CC -1- FUNCTION: REGULATES THE PRODUCTION OF A SIGNAL THAT INDUCES  
CC HAIR (TRICHOME) PRECURSOR CELLS ON LEAF PRIMORDIA TO  
CC DIFFERENTIATE.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- SIMILARITY: Contains 2 Myb-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M79448; AAC97387.1; -.  
DR EMBL; AB006078; BAA86879.1; -.  
DR EMBL; AF263690; AAL01215.1; -.  
DR EMBL; AF263693; AAL01218.1; -.  
DR EMBL; AF263694; AAL01219.1; -.  
DR EMBL; AF263695; AAL01220.1; -.  
DR EMBL; AF263696; AAL01221.1; -.  
DR EMBL; AF263697; AAL01222.1; -.  
DR EMBL; AP000371; BAB02538.1; -.  
DR HSSP; P01103; IPOM.  
DR TRANSFAC; T01588; -.  
DR InterPro; IPR001005; Myb DNA binding.  
DR Pfam; PF00249; myb DNA-binding; 2.  
DR SMART; SM00717; SANT; 2.  
DR PROSITE; PS00037; MYB\_1; 1.  
DR PROSITE; PS00334; MYB\_2; 1.  
DR PROSITE; PS50090; MYB\_3; 2.  
KM Nuclear protein; DNA-binding; Repeat; Transcription regulation.  
FT DNA\_BIND 11 63 MYB 1.  
FT DNA\_BIND 64 114 MYB 2.  
FT CONFLICT 149 149 C->R (IN REF. 1).  
FT CONFLICT 222 224 GDV->SDI (IN REF. 1).  
SQ SEQUENCE 228 AA; 2638 MW; 822A8646F46F6B9 CRC64;  
Query March 5.1%; Score 14; DB 1; Length 228;  
Best Local Similarity 100.0%; Pred. No. 8.2e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 79 RLHKLGNRMSLIA 92  
Db 82 RLHKLGNRMSLIA 95  
RESULT 5  
MYB1\_HORVU STANDARD; PRT; 267 AA.  
ID MYB1\_HORVU  
AC P20026;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myb-related protein Hv1.  
GN MYB1.  
OS Hordeum vulgare (Barley).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Hordeum.  
OC NCB1\_TaxID=4513;  
OX NCB1 [1]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Abyssinian 2231;  
RX MEDLINE=89313655; PubMed=2664447;  
RA Marocco A., Wissembach M., Becker D., Paz-Ares J., Seidler H.,  
RA Salami F., Rohde W.;  
RT "Multiple genes are transcribed in Hordeum vulgare and Zea mays that  
RT carry the DNA binding domain of the myb oncoproteins.";

RL Mol. Gen. Genet. 216:183-187(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Abyssinian 2231; TISSUE=Leaf;  
RX MEDLINE=94035190; PubMed=8220488;  
RA Wissembach M., Ueberlacker B., Vogt F., Becker D., Salami F.,  
RA Rohde W.;  
RT "Myb genes from Hordeum vulgare, tissue-specific expression of  
RT chimeric Myb promoter/Gus genes in transgenic tobacco";  
RL Plant J. 4:411-422(1993).  
CC -1- FUNCTION: POSSIBLE TRANSCRIPTION ACTIVATOR IN RESPONSE TO AN  
CC EXTERNAL SIGNAL. MAY BE INVOLVED IN THE REGULATION OF FLAVONOID  
CC BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- TISSUE SPECIFICITY: GERMINATING SEED AND APICAL MERISTEM OF SHOOT  
CC AND ROOT.  
CC -1- SIMILARITY: Contains 2 Myb-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X70877; CAAS0222.1; -.  
DR EMBL; X70879; CAAS0224.1; -.  
DR PTR; S61506; S61506.  
DR HSSP; P01103; IPOM.  
DR TRANSFAC; T02887; -.  
DR InterPro; IPR001005; Myb DNA binding.  
DR Pfam; PF00249; myb DNA-binding; 2.  
DR SMART; SM00717; SANT; 2.  
DR PROSITE; PS00037; MYB\_1; 1.  
DR PROSITE; PS00334; MYB\_2; 1.  
DR PROSITE; PS50090; MYB\_3; 2.  
KM Nuclear protein; DNA-binding; Activator; Transcription regulation;  
FT DNA\_BIND 9 61 MYB 1.  
FT DNA\_BIND 62 112 MYB 2.  
SQ SEQUENCE 267 AA; 29740 MW; DBCCDCBBA847749 CRC64;  
Query March 4.4%; Score 12; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.00011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 88 WSLIAGRLPGR 99  
Db 89 WSLIAGRLPGR 100  
RESULT 6  
MYB1\_MAIZE STANDARD; PRT; 399 AA.  
ID MYB1\_MAIZE  
AC P27898; P27899;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Myb-related protein P.  
GN P.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
OC NCB1\_TaxID=4577;  
OX NCB1 [1]  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91271238; PubMed=2052542;  
RA Grotenwald E., Ahma P., Peterson T.;  
RT "Alternatively spliced products of the maize P gene encode proteins  
RT with homology to the DNA-binding domain of myb-like transcription  
RT factors.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 88:4587-4591(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92275319; PubMed=1317315;
RA Athma P., Grotewold E., Peterson T.;
RT "insertional mutagenesis of the maize P gene by intragenic
RT transposition of Ac.";
RL Genetics 131:199-209(1992).
CC -1- FUNCTION: TRANSCRIPTION FACTOR POSTULATED TO REGULATE THE
CC BIOSYNTHETIC PATHWAY OF A FLAVONOID-DERIVED PIGMENT IN CERTAIN
CC FLOREL TISSUES.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P27898-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P27898-2; Sequence=VSP_003301, VSP_003302;
CC -1- SIMILARITY: Contains 2 Myb-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, WJ3028; AAA33500.1; -
DR EMBL, WJ3029; AAA33501.1; -
DR EMBL, Z11879; CAA77939.1; -
DR PIR, A39697; A39697.
DR PIR, B39697; B39697.
DR HSSP, P01103; 1POM.
DR TRANSFAC, T01590; -
DR TRANSFAC, T01591; -
DR MaizEDB; 69180; -
DR MaizEDB; 69181; -
DR InterPro; IPR001005; Myb DNA binding.
DR Pfam; PF00249; myb DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS50090; MYB_3; 2.
DR Nuclear protein; DNA-binding; Repeat; Transcription regulation;
KM Alternative splicing.
FT DNA_BIND 9 61 MYB 1.
FT DNA_BIND 62 112 MYB 2.
FT DOMAIN 178 181 POLY-ALA.
FT DOMAIN 305 311 POLY-LEU.
FT VARSPPLIC 89 152 WSILASHLPGRDNEIKYWNHLSROIHYRRKRYAGPDD
TATAIDMSKQSDRRRGSTPG -> RHIMTADISPEST
VRCILPRGALVYLTLPQSPQRTARTYDRGSLALNSVRC
FRCPVPSRW (in isoform Short).
FT FT FT /FTId=VSP_003301.
FT FT FT /FTId=VSP_003302.
FT FT FT Missing (in isoform Short).
SQ SEQUENCE 399 AA; 43756 MW; EB025B00A44CF5D0 CRC64;
Query Match 3.3%; Score 9; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 84 LGNRWSLIA 92
DB 85 LGNRWSLIA 93
RESULT 7
M3K4_MOUSE STANDARD; PRT; 1597 AA.
ID M3K4_MOUSE 008648; 008649; 070124;
AC 30-MAY-2000 (Rel. 39, Created)

```

```

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 4 (EC 2.7.1.-)
DE (MAPK/ERK kinase kinase 4) (MEKK 4).
GN MAP3K4 OR MEKK4.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RC TISSUE=Brain;
RX MEDLINE=97236778; PubMed=9079650;
RA Gervins P., Blank J.L., Johnson G.L.;
RT "Cloning of a novel mitogen-activated protein kinase kinase,
RT MEKK4, that selectively regulates the c-Jun amino terminal kinase
RT pathway.";
RL J. Biol. Chem. 272:8288-8295(1997).
RN [2]
RP SEQUENCE OF 363-1049 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Ectoplacental cone;
RX MEDLINE=97422605; PubMed=9268631;
RA Schweifer N., Valk P.J., Delwel R., Cox R., Francis F.,
RA Meier-Ewert S., Lehrach H., Barlow D.P.;
RT "Characterization of the C3 YAC contig from proximal mouse chromosome
RT 17 and analysis of allelic expression of genes flanking the imprinted
RT Igf2r gene.";
RL Genomics 43:285-297(1997).
CC -1- FUNCTION: ACTIVATES THE CSBP2, P38 AND JNK MAPK PATHWAYS, BUT NOT
CC -1- THE ERK PATHWAY. SPECIFICALLY PHOSPHORYLATES AND ACTIVATES MAP2K4
CC AND MAP2K6.
CC -1- SUBCELLULAR LOCATION: LOCALIZED IN PERINUCLEAR VESICULAR-LIKE
CC STRUCTURES. PROBABLY GOLGI-ASSOCIATED VESICLES.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=A;
CC IsoId=O08648-1; Sequence=Displayed;
CC Name=B;
CC IsoId=O08648-2; Sequence=VSP_004885.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. HIGH EXPRESSION WAS FOUND IN
CC SKELETAL MUSCLE, KIDNEY, TESTIS FOLLOWED BY HEART BRAIN AND LUNG.
CC LOW EXPRESSION WAS FOUND IN SPLEEN.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MAP
CC KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, U85607; AAC53126.1; -
DR EMBL, U85608; AAC53127.1; -
DR EMBL, U66240; AAC08286.1; -
DR HSSP, P24941; 1HCL.
DR MGD; MG1:1346875; Map3k4.
DR InterPro; IPR000719; Prot Kinase.
DR InterPro; IPR002290; Ser Thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR Prodom; PD0000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KM Alternative splicing.
FT DOMAIN 1178 1182 POLY-ALA.
FT DOMAIN 1332 1590 PROTEIN_KINASE.
FT NE_BIND 1338 1346 ATP (BY SIMILARITY).
FT BINDING 1361 1361 ATP (BY SIMILARITY).
FT ACT_SITE 1452 1452 BY SIMILARITY.

```



```

FT  VARSPLIC  1162  1213  Missing (in isoform B).
FT  CONFLICT  363  364  /FTID=VSP_004885.
FT  CONFLICT  473  473  SL -> NS (IN REF. 2).
SQ  SEQUENCE  1597 AA; 179948 MW; E84AEAAE92D103A4 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  76 LIIRLHKL 83
Db  533 LIIRLHKL 540

RESULT 8
M3K4_HUMAN  STANDARD;  PRT;  1607 AA.
ID  M3K4_HUMAN
AC  Q9Y6R4; Q92612;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Mitogen-activated protein kinase kinase 4 (MEKK 4) (MAP three kinase
DE  1).
GN  MAP3K4 OR MAP3KK4 OR MEKK4 OR MTK1 OR KIAA0213.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND MUTAGENESIS OF LYS-1371.
RC  TISSUE=Fetal liver, and skeletal muscle;
RA  MEDLINE=97449143; PubMed=9305639;
RA  Takekawa M., Posas F., Saito H.;
RT  "A human homolog of the yeast Sak2/Spk2 MAP kinase kinases,
RT  MTK1, mediates stress-induced activation of the p38 and JNK
RT  pathways."
RL  EMBO J. 16:4973-4982 (1997).
RN  [2]
RP  SEQUENCE OF 68-1607 FROM N.A. (ISOFORM B).
RC  TISSUE=Bone marrow;
RA  MEDLINE=97191544; PubMed=9039502;
RA  Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayashi Y.,
RA  Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT  "Prediction of the coding sequences of unidentified human genes. VI.
RT  The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT  analysis of cDNA clones from cell line KG-1 and brain."
RL  DNA Res. 3:321-329 (1996).
CC  -1- FUNCTION: ACTIVATES THE GSBP2, P38 AND JNK MAPK PATHWAYS, BUT NOT
CC  THE ERK PATHWAY. SPECIFICALLY PHOSPHORYLATES AND ACTIVATES MAP2K4
CC  AND MAP2K6.
CC  -1- ALTERNATIVE PRODUCTS:
CC  Event=Alternative splicing; Named Isoforms=2;
CC  Name=A;
CC  IsoId=Q9Y6R4-1; Sequence=Displayed;
CC  Name=B;
CC  IsoId=Q9Y6R4-2; Sequence=VSP_004884;
CC  -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN HEART, PLACENTA,
CC  SKELETAL MUSCLE AND PANCREAS, AND AT LOWER LEVELS IN OTHER
CC  TISSUES.
CC  -1- DOMAIN: THE N-TERMINAL NON-CATALYTIC DOMAIN (1-606) CONTAINS AN
CC  ESSENTIAL DOMAIN FOR THE DOMINANT-INHIBITORY EFFECT.
CC  -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MAP
CC  KINASE KINASE SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.ebi.ac.uk/announce/
CC  or send an email to license@ebi.ac.uk).

```

```

CC  -----
CC  EMBL; AF002715; AAB68804.1; -
DR  EMBL; D66968; BAA13204.1; -
DR  PIR; T03022; T03022.
DR  HSSP; P24941; 1HCL.
DR  Genew; HGNC:6856; MAP3K4.
DR  MIM; 602425; -
DR  GO; GO:0004709; F:MAP kinase kinase kinase activity; TAS.
DR  GO; GO:0000186; P:activation of MAPKK; TAS.
DR  GO; GO:0007254; P:JNK cascade; TAS.
DR  GO; GO:0006950; P:response to stress; TAS.
DR  InterPro; IPR000719; Prot_kinase.
DR  InterPro; IPR002290; Ser_thr_kinase.
DR  Pfam; PF00069; pkinase; 1.
DR  Prodom; PD000001; Prot_kinase; 1.
DR  SMART; SM00220; S_TKc_1.
DR  PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR  PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR  PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KM  Transferase; Serine/threonine-protein kinase; ATP-binding;
KW  Alternative splicing.
FT  DOMAIN  4 7
FT  DOMAIN  25 38
FT  DOMAIN  1190 1201
FT  NP_BIND  1342 1600
FT  NP_BIND  1348 1356
FT  BINDING  1371 1371
FT  ACT_SITE  1462 1462
FT  VARSPLIC  1175 1223
FT  MUTAGEN  1371 1371
FT  CONFLICT  791 791
SQ  SEQUENCE  1607 AA; 181550 MW; 24BC4597730B5287 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  76 LIIRLHKL 83
Db  542 LIIRLHKL 549

RESULT 9
NARP_HAEM  STANDARD;  PRT;  208 AA.
ID  NARP_HAEM
AC  P44845;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Nitrate/nitrite response regulator protein homolog.
GN  NARP OR HI0726.
OS  Haemophilus influenzae.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC  Pasteurellaceae; Haemophilus.
OX  NCBI_TaxID=127;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Rd / KM20 / ATCC 51907;
RX  MEDLINE=95350630; PubMed=7542800;
RA  Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA  Kesteven K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA  Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA  Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA  Letererack T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA  Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
RA  Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA  Venter J.C.;
RT  "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT  Rd."
RL  Science 269:496-512 (1995).
CC  -1- FUNCTION: COULD ACTIVATE THE EXPRESSION OF A FORMATE DEHYDROGENASE

```



```

CC OPERONS AND COULD REPRESS THE TRANSCRIPTION OF THE FUMARATE
CC REDUCTASE (FRDABCD) OPERON.
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC -1- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U33756; AAC22384.1; -.
CC PIR: A64089; A64089.
CC HSSP: P10957; IRNL.
CC TIGR: H10726; -.
CC InterPro: IPR000792; HTH_LuxR.
CC InterPro: IPR001789; Response_reg.
CC Pfam: PF00196; Gers. 1.
CC Pfam: PF00072; Response_reg. 1.
CC PRINTS: PR00038; HTHLUXR.
CC ProDom: PD000307; HTH_LuxR; 1.
CC ProDom: PD000039; Response_reg; 1.
CC SMART: SM00421; HTH_LuxR; 1.
CC SMART: SM00448; REC; 1.
CC PROSITE: PS00622; HTH_LUXR_FAMILY; 1.
CC PROSITE: PS50110; RESPONSE_REGULATORY; 1.
CC Transcription regulation; Activator; Repressor; DNA-binding;
CC ATP-binding; Phosphorylation; Sensory transduction; Complete proteome.
CC DOMAIN 6 122 RESPONSE_REGULATORY.
CC MOD RES 57 57 PHOSPHORYLATION (BY SIMILARITY).
CC FT DNA_BIND 166 185 H-T-H MOTIF (BY SIMILARITY).
CC SEQUENCE 208 AA; 23078 MW; 0B72119BEC4906C7 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 NLIRKLN 116
Db 186 NLIRKLN 192

RESULT 10
NARP_ECOLI
ID NARP_ECOLI STANDARD; PRT; 215 AA.
AC P31802;
DT 01-JUL-1993 (Rel. 26, Last Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nitrate/nitrite response regulator protein narp.
GN NARP OR B2193.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93273691; PubMed=8501030;
RA Rabin R.S., Stewart V.;
RT "Dual response regulators (NarX and Narp) interact with dual sensors
RT expression in Escherichia coli K-12."
RL J. Bacteriol. 175:3259-3268(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / BHE2600;
RA Richerich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
RA Church G.M.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.

```

```

RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glaeser J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: THIS PROTEIN ACTIVATES THE EXPRESSION OF THE NITRATE
CC REDUCTASE (NARXHI) AND FORMATE DEHYDROGENASE-N (FONGHI) OPERONS
CC AND REPRESSSES THE TRANSCRIPTION OF THE FUMARATE REDUCTASE
CC (FRDABCD) OPERON IN RESPONSE TO A NITRATE/NITRITE INDUCTION SIGNAL
CC TRANSMITTED BY EITHER THE NARX OR NARO PROTEINS.
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC -1- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS. SIMILAR TO THE NARL PROTEIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U11273; AAA24200.1; -.
CC EMBL: U00008; AAA16411.1; ALT_INIT.
CC EMBL: AE000309; AAC75253.1; -.
CC PIR: A40584; A40584.
CC HSSP: P10957; IRNL.
CC EcoGene: EG11527; narp.
CC InterPro: IPR000792; HTH_LuxR.
CC InterPro: IPR001789; Response_reg.
CC Pfam: PF00196; Gers. 1.
CC Pfam: PF00072; Response_reg. 1.
CC PRINTS: PR00038; HTHLUXR.
CC ProDom: PD000307; HTH_LuxR; 1.
CC ProDom: PD000039; Response_reg; 1.
CC SMART: SM00421; HTH_LuxR; 1.
CC SMART: SM00448; REC; 1.
CC PROSITE: PS00622; HTH_LUXR_FAMILY; 1.
CC PROSITE: PS50110; RESPONSE_REGULATORY; 1.
CC Transcription regulation; Activator; Repressor; DNA-binding;
CC ATP-binding; Phosphorylation; Sensory transduction;
CC Nitrate assimilation; Complete proteome.
CC DOMAIN 8 124 RESPONSE_REGULATORY.
CC MOD RES 59 59 PHOSPHORYLATION (BY SIMILARITY).
CC FT DNA_BIND 171 190 H-T-H MOTIF (BY SIMILARITY).
CC SEQUENCE 215 AA; 23575 MW; C7121668E3130956 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 NLIRKLN 116
Db 191 NLIRKLN 197

RESULT 11
PSB2_CRYNE
ID PSB2_CRYNE STANDARD; PRT; 224 AA.
AC Q00826;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable proteasome subunit beta type 2 (EC 3.4.25.1).
GN CPRL.
OS Cryococcus neoformans (Filobasidiella neoformans).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.

```

```

OX NCBI_TaxID=5207;
RN SEQUENCE FROM N.A.
RX MEDLINE=96239006; PubMed=8675296;
RA Chang Y.C., Penoyer L.A., Kwon-Chung K.J.;
RT "The second capsule gene of Cryptococcus neoformans, CAP64, is
  essential for virulence."
RL Infect. Immun. 64:1977-1983 (1996).
CC -1- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
  WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH ARG,
  PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT NEUTRAL OR
  SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT PROTEOLYTIC
  ACTIVITY (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Cleavage at peptide bonds with very broad
  specificity.
CC -1- PATHWAY: Involved in an ATP/ubiquitin-dependent non-lysosomal
  proteolytic pathway.
CC -1- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
  SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE (BY
  SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
CC -1- SIMILARITY: Belongs to peptidase family T1B.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL: L40028; AAB06582.1; -.
CC HSSP: P22141; IRYP.
DR MEROPS: T01.984; -.
DR InterPro: IPR000243; Proteasome_B.
DR InterPro: IPR001353; Protease_protease.
DR Pfam: PF00227; Proteasome; 1.
DR PROSITE: PS00854; PROTEASOME_B; FALSE_NEG.
KM Proteasome; Hydrolase; Protease.
SQ SEQUENCE 224 AA; 25239 MW; ABEEF03E76170886 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 224;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 LIRKCID 29
DB 160 LIRKCID 166

RESULT 12
MYB3_MAIZE STANDARD; PRT; 255 AA.
AC P20025;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myb-related protein Zm38.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89313655; PubMed=2664447;
RA Marocco A., Wissenbach M., Becker D., Paz-Ares J., Saedler H.,
RA Salamini F., Rohde W.;
RT "Multiple genes are transcribed in Hordeum vulgare and Zea mays that
  carry the DNA binding domain of the myb oncoproteins."
RL Mol. Genet. 216:183-187 (1989).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 2 Myb-like domains.

```

```

DR PIR; S04899; S04899.
DR HSSP; P01103; 1POM.
DR TRANSFAC; T02959; -.
DR MaizeDB; 69594; -.
DR InterPro: IPR001005; Myb_DNA_binding.
DR Pfam; PF00249; myb_DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS50090; MYB_3; 2.
KM Nuclear protein; DNA-binding; Repeat; Transcription regulation.
FT DNA BIND 9 61 MYB 1.
FT DNA BIND 62 112 MYB 2.
SQ SEQUENCE 255 AA; 27568 MW; 26AADC8B0633E06 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 KSCRLRW 56
DB 51 KSCRLRW 57

RESULT 13
MYB3_HORVU STANDARD; PRT; 302 AA.
AC P20037;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myb-related protein Hv33.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Abyssinian 2231; TISSUE=leaf;
RX MEDLINE=9405190; PubMed=8220488;
RA Wissenbach M., Ueberlackner B., Vogt F., Becker D., Salamini F.,
RA Rohde W.;
RT "Myb genes from Hordeum vulgare: tissue-specific expression of
  chimeric Myb promoter/Gus genes in transgenic tobacco."
RL Plant J. 4:411-422 (1993).
RN [2]
RP SEQUENCE OF 35-302 FROM N.A.
RX STRAIN=cv. Abyssinian 2231;
RX MEDLINE=89313655; PubMed=2664447;
RA Marocco A., Wissenbach M., Becker D., Paz-Ares J., Saedler H.,
RA Salamini F., Rohde W.;
RT "Multiple genes are transcribed in Hordeum vulgare and Zea mays that
  carry the DNA binding domain of the myb oncoproteins."
RL Mol. Genet. 216:183-187 (1989).
CC -1- FUNCTION: POSSIBLE TRANSCRIPTION ACTIVATOR IN RESPONSE TO AN
  EXTERNAL SIGNAL. MAY BE INVOLVED IN THE REGULATION OF FLAVONOID
  BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: GERMINATING SEED AND APICAL MERISTEM OF SHOOT
  AND ROOT.
CC -1- SIMILARITY: Contains 2 Myb-like domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL; X70881; CAA50226.1; -.

```

```

DR EMBL, X70878, CAA50223.1; -.
DR PIR, S31818, S31818.
DR HSSP, P01103, IPOM.
DR TRANSFAC: T02889; -.
DR InterPro: IPR001005, Myb DNA-binding.
DR Pfam, PF00249, myb-DNA-binding; 2.
DR SMART, SM00717, SANT; 2.
DR PROSITE, PS00037, MYB_1; 1.
DR PROSITE, PS00334, MYB_2; 1.
DR PROSITE, PS50090, MYB_3; 2.
KM Nuclear protein, DNA-binding; Activator; Transcription regulation;
KW Repeat.
FT DNA_BIND 11 63 MYB 1.
FT DNA_BIND 64 114 MYB 2.
FT CONFLICT 259 259 A->P (IN REF. 2).
SQ SEQUENCE 302 AA; 33003 MW; 89017418C3631599 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 KSCRLRW 56
Db 53 KSCRLRW 59

RESULT 14
YACB_BACSU STANDARD; PRT; 315 AA.
ID YACB_BACSU STANDARD; PRT; 315 AA.
AC E37526;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypochemical protein yacB.
GN YACB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RL subtilis chromosome containing the replication origin.";
DNA Res. 1:1-14(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Besieres P., Bolotin A., Borchert S.,
RA Boursier R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gilm S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasaahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogiwara A., Oudega B., Park S.H.,
RA Pardo V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Priesen E., Pujic P., Purnelle B., Rapoport G., Rev M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadele Y.,
RA Sato T., Scallan E., Schleich S., Schroeder R., Scofield F.,
RA Setiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

```

```

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein E.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D26185, BAA05244.1; -.
DR EMBL, Z99104, CAB11784.1; -.
DR PIR, S66038, S66038.
DR Subtilist; BG10072; YACB.
KW Hypochemical protein, Complete proteome.
SQ SEQUENCE 315 AA; 37978 MW; E9BE25481FA13519 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 204 EEVINY 210
Db 183 EEVINY 189

RESULT 15
MDHC_HUMAN STANDARD; PRT; 333 AA.
ID MDHC_HUMAN STANDARD; PRT; 333 AA.
AC P40925;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Malate dehydrogenase, cytoplasmic (EC 1.1.1.37).
GN MDH1 OR MDHA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=96230334; PubMed=8766100;
RA Tanaka T., Inazawa J., Nakamura Y.;
RT "Molecular cloning and mapping of a human cDNA for cytosolic malate
RL dehydrogenase (MDH1).";
Genomics 32:128-130(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Lo A.S.Y., Waye M.M.Y.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Yoshizaki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettlemen M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shurchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywiński M.I., Skalska U., Smallus D.E.,  
RA Smerich A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [4]  
RP SEQUENCE OF 167-180.  
RC TISSUE=Heart;  
RX MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
RT 1994."  
RL Electrophoresis 15:1459-1465 (1994).  
CC -1- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.  
CC -1- SUBUNIT: Homodimer.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch)  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
CC  
CC EMBL; D55654; BAA09513.1; -.  
DR EMBL; U20352; AAC16436.1; -.  
DR EMBL; BC001484; AAH01484.1; -.  
DR HSSP; P11708; 4MDH.  
DR HSC-2DPAGE; P40925; HUMAN.  
DR Genew; HGNC:6970; MDH1.  
DR MIM; 154200; -.  
DR GO; GO:0005829; C:cytosol; TAS.  
DR GO; GO:0004470; E:malic enzyme activity; TAS.  
DR InterPro; IPR001236; 1dh.  
DR InterPro; IPR001252; mdh.  
DR Pfam; PF00056; 1dh; 1.  
DR Pfam; PF02866; 1dh; 1.  
DR Prodom; PD003052; Mdh; 1.  
DR PROSITE; PS00068; MDH; 1.  
KM Oxidoreductase; Tricarboxylic acid cycle; NAD.  
FT INIT MET 0  
FT ACT SITE 158 158 PROTON-RELAY.  
FT BINDING 161 161 SUBSTRATE CARBOXYL.  
FT ACT SITE 186 186 PROTON-RELAY.  
SQ SEQUENCE 333 AA; 36295 MW; 886F9A79FE5E888 CRC64;

Query Match	2.6%	Score 7	DB 1	Length 333
Best Local Similarity	100.0%	Pred. No. 23		
Matches 7; Conservative	0	Mismatches 0	Indels 0	Gaps 0

QY	99	TANDVKN	105
Db	173	TANDVKN	179

Search completed: January 29, 2004, 20:20:06  
Job time : 38 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: January 29, 2004, 20:17:07 ; Search time 41 seconds  
(without alignments)  
1724.547 Million cell updates/sec

Title: US-10-033-190-2  
Perfect score: 274  
Sequence: 1 MNSTMSISLGVKSGSWTDE.....QQGVQNDPSAIDLNLLD 274

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rhodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	38	13.9	120	10	Q9M714
2	38	13.9	127	10	Q9M717
3	38	13.9	255	10	Q9M721
4	38	13.9	255	10	Q9M720
5	34	12.4	137	10	Q9M716
6	34	12.4	139	10	Q9M718
7	32	11.7	139	10	Q9M718
8	32	11.7	246	10	Q9M719
9	32	11.7	249	10	Q9M719
10	28	10.2	255	10	Q9M719
11	27	9.9	212	10	Q9M719
12	27	9.9	248	10	Q9M719
13	25	9.1	179	10	Q9M719
14	25	9.1	181	10	Q9M719
15	25	9.1	184	10	Q9M719
16	25	9.1	193	10	Q9M719

17	25	9.1	193	10	Q49018	Q49018 gossypium h
18	25	9.1	193	10	Q49FG2	Q49FG2 gossypium h
19	25	9.1	247	10	Q8L5P2	Q8L5P2 vicis labru
20	25	9.1	250	10	Q8L5P3	Q8L5P3 vicis labru
21	25	9.1	307	10	Q8L5P1	Q8L5P1 vicis labru
22	21	7.7	113	10	Q9J3V9	Q9J3V9 zea luxuria
23	21	7.7	113	10	Q94F21	Q94F21 zea luxuria
24	21	7.7	118	10	Q8S4Z5	Q8S4Z5 zea maye (m
25	21	7.7	157	10	Q8H267	Q8H267 gossypium h
26	21	7.7	157	10	Q8H266	Q8H266 gossypium h
27	21	7.7	157	10	Q8H265	Q8H265 gossypium h
28	21	7.7	157	10	Q8H264	Q8H264 gossypium h
29	21	7.7	157	10	Q8H263	Q8H263 gossypium h
30	21	7.7	187	10	Q94EN3	Q94EN3 fragaria an
31	21	7.7	198	10	Q49017	Q49017 gossypium h
32	21	7.7	215	10	Q9F212	Q9F212 bolanum tub
33	21	7.7	254	10	Q9ATD2	Q9ATD2 gossypium h
34	21	7.7	256	10	Q22451	Q22451 zea maye (m
35	21	7.7	264	10	Q41869	Q41869 zea maye (m
36	21	7.7	266	10	Q41868	Q41868 zea maye (m
37	21	7.7	267	10	Q22450	Q22450 zea maye (m
38	21	7.7	271	10	Q41842	Q41842 zea maye (m
39	21	7.7	271	10	Q8GRW4	Q8GRW4 zea maye (m
40	20	7.3	45	10	Q8L6R5	Q8L6R5 gossypium h
41	20	7.3	45	10	Q8L6R3	Q8L6R3 gossypium h
42	20	7.3	46	10	Q8L6P6	Q8L6P6 gossypium h
43	20	7.3	46	10	Q8L6M1	Q8L6M1 gossypium h
44	20	7.3	46	10	Q8L6M9	Q8L6M9 gossypium h
45	20	7.3	46	10	Q8L6R6	Q8L6R6 gossypium h
46	20	7.3	302	10	Q9ATD5	Q9ATD5 gossypium h
47	19	6.9	130	10	Q9ZTC4	Q9ZTC4 arabidopsis
48	19	6.9	201	10	Q9SQJ6	Q9SQJ6 arabidopsis
49	19	6.9	201	10	Q9LTF7	Q9LTF7 arabidopsis
50	18	6.6	45	10	Q49804	Q49804 arabidopsis
51	18	6.6	129	10	Q9M715	Q9M715 petunia axi
52	17	6.2	44	10	Q9SQJ2	Q9SQJ2 zea maye (m
53	17	6.2	46	10	Q8L6R7	Q8L6R7 gossypium h
54	17	6.2	46	10	Q8L6M6	Q8L6M6 gossypium h
55	17	6.2	46	10	Q8L6P8	Q8L6P8 gossypium h
56	17	6.2	60	10	Q8L5P0	Q8L5P0 vicis labru
57	17	6.2	96	10	Q9AQY0	Q9AQY0 zea perenul
58	17	6.2	96	10	Q94FY9	Q94FY9 tripasacum d
59	17	6.2	96	10	Q9ATY1	Q9ATY1 zea diplope
60	17	6.2	96	10	Q9AQY1	Q9AQY1 zea diplope
61	17	6.2	96	10	Q9ATY4	Q9ATY4 zea diplope
62	17	6.2	113	10	Q94FZ2	Q94FZ2 zea maye (s
63	17	6.2	113	10	Q9JWL7	Q9JWL7 zea luxuria
64	17	6.2	113	10	Q9JWL8	Q9JWL8 zea maye (s
65	17	6.2	166	10	Q8S3X8	Q8S3X8 oryza sativ
66	17	6.2	179	10	Q8S3Z6	Q8S3Z6 zea maye (m
67	17	6.2	195	10	Q9C9A5	Q9C9A5 arabidopsis
68	17	6.2	209	10	Q8H254	Q8H254 gossypium h
69	17	6.2	213	10	Q39156	Q39156 arabidopsis
70	17	6.2	218	10	Q8H253	Q8H253 gossypioide
71	17	6.2	226	10	Q49021	Q49021 gossypium h
72	17	6.2	226	10	Q8H256	Q8H256 gossypium h
73	17	6.2	226	10	Q8H255	Q8H255 gossypium r
74	17	6.2	271	10	Q9ATD3	Q9ATD3 gossypium h
75	17	6.2	272	10	Q9ZRS0	Q9ZRS0 oryza sativ
76	17	6.2	275	10	Q23892	Q23892 oryza sativ
77	17	6.2	321	10	Q23891	Q23891 oryza sativ
78	16	5.8	35	10	Q8GUX3	Q8GUX3 picea maria
79	16	5.8	42	10	Q9XIU3	Q9XIU3 glycine max
80	16	5.8	42	10	Q9XIS9	Q9XIS9 glycine max
81	16	5.8	42	10	Q9XIT2	Q9XIT2 glycine max
82	16	5.8	45	10	Q49798	Q49798 arabidopsis
83	16	5.8	45	10	Q8L6Q3	Q8L6Q3 gossypium h
84	16	5.8	45	10	Q8L6Q1	Q8L6Q1 gossypium h
85	16	5.8	45	10	Q8L6P9	Q8L6P9 gossypium h
86	16	5.8	45	10	Q8L6R9	Q8L6R9 gossypium h
87	16	5.8	45	10	Q8L6S4	Q8L6S4 gossypium h
88	16	5.8	46	10	Q8L6L7	Q8L6L7 gossypium h
89	16	5.8	46	10	Q8L6Q2	Q8L6Q2 gossypium h

90	16	5.8	46	10	08L6N8	08L6N8	gossypium h	163	13	4.7	376	10	09SBF7	09SBF7	arabidopsis
91	16	5.8	46	10	08L6P1	08L6P1	gossypium h	164	13	4.7	388	10	04O920	04O920	picea maria
92	16	5.8	46	10	08L6L6	08L6L6	gossypium h	165	12	4.4	35	10	08GXU0	08GXU0	gossypium h
93	16	5.8	46	10	08L6R8	08L6R8	gossypium h	166	12	4.4	43	10	08S0E4	08S0E4	zea mays (m
94	16	5.8	46	10	08L6N6	08L6N6	gossypium h	167	12	4.4	43	10	09S0Q9	09S0Q9	zea mays (m
95	16	5.8	113	10	094F20	094F20	zea luxurita	168	12	4.4	43	10	09S0Q2	09S0Q2	zea mays (m
96	14	5.1	128	10	092XD2	092XD2	zea luxurita	169	12	4.4	43	10	09S0Q3	09S0Q3	zea mays (m
97	14	5.1	148	10	092ZD2	092ZD2	perilla fru	170	12	4.4	43	10	09S0A9	09S0A9	zea mays (m
98	14	5.1	181	10	09LFB2	09LFB2	arabidopsis	171	12	4.4	43	10	09S0E1	09S0E1	zea mays (m
99	14	5.1	201	10	039I09	039I09	arabidopsis	172	12	4.4	43	10	09S0E1	09S0E1	zea mays (m
100	14	5.1	203	10	09S8I0	09S8I0	arabidopsis	173	12	4.4	45	10	08L6P4	08L6P4	arabidopsis
101	14	5.1	219	10	096276	096276	arabidopsis	174	12	4.4	45	10	049792	049792	arabidopsis
102	14	5.1	223	10	0947R5	0947R5	arabidopsis	175	12	4.4	46	10	08L6P3	08L6P3	gossypium h
103	14	5.1	228	10	093V34	093V34	arabidopsis	176	12	4.4	60	10	08L5N9	08L5N9	vitis labru
104	14	5.1	228	10	093W30	093W30	arabidopsis	177	12	4.4	183	10	08GV56	08GV56	zea mays (m
105	14	5.1	228	10	093W40	093W40	arabidopsis	178	12	4.4	188	10	08S444	08S444	sorghum bic
106	14	5.1	228	10	0947R6	0947R6	arabidopsis	179	12	4.4	222	10	P81393	P81393	antirrhinum
107	14	5.1	228	10	0947R6	0947R6	arabidopsis	180	12	4.4	226	10	038851	038851	arabidopsis
108	14	5.1	228	10	0947R6	0947R6	arabidopsis	181	12	4.4	246	10	09S7Y2	09S7Y2	arabidopsis
109	13	4.7	96	10	09AQY2	09AQY2	zea diptope	182	12	4.4	247	10	09ATD9	09ATD9	arabidopsis
110	13	4.7	96	10	09AQY2	09AQY2	zea diptope	183	12	4.4	249	10	038850	038850	arabidopsis
111	13	4.7	96	10	09AQY3	09AQY3	zea diptope	184	12	4.4	257	10	09S9K9	09S9K9	arabidopsis
112	13	4.7	96	10	09AQY3	09AQY3	zea diptope	185	12	4.4	253	10	08H268	08H268	gossypioide
113	13	4.7	104	10	0947R4	0947R4	arabidopsis	186	12	4.4	264	10	09ATD1	09ATD1	gossypium h
114	13	4.7	157	10	08S321	08S321	sorghum bic	187	12	4.4	265	10	08H269	08H269	gossypium h
115	13	4.7	165	10	08S3Y9	08S3Y9	sorghum bic	188	12	4.4	268	10	08H270	08H270	gossypium r
116	13	4.7	184	10	08S3Y6	08S3Y6	sorghum bic	189	12	4.4	273	10	043525	043525	lycopersico
117	13	4.7	198	10	09AVB7	09AVB7	lilium hybr	190	12	4.4	276	10	023890	023890	oryza sativ
118	13	4.7	203	10	P81396	P81396	antirrhinum	191	12	4.4	282	10	09S2P1	09S2P1	arabidopsis
119	13	4.7	205	10	08S320	08S320	sorghum bic	192	12	4.4	282	10	049774	049774	arabidopsis
120	13	4.7	206	10	P81391	P81391	antirrhinum	193	12	4.4	294	10	043436	043436	gossypium h
121	13	4.7	206	10	08L8W8	08L8W8	arabidopsis	194	12	4.4	294	10	08H271	08H271	gossypium h
122	13	4.7	206	10	09SSA1	09SSA1	arabidopsis	195	11	4.0	45	10	049766	049766	arabidopsis
123	13	4.7	214	10	08L8C0	08L8C0	arabidopsis	196	11	4.0	45	10	09S8M3	09S8M3	arabidopsis
124	13	4.7	214	10	09S8P9	09S8P9	arabidopsis	197	11	4.0	46	10	08L6R3	08L6R3	gossypium h
125	13	4.7	216	10	09FMW0	09FMW0	arabidopsis	198	11	4.0	267	10	09M9U2	09M9U2	arabidopsis
126	13	4.7	217	10	P93474	P93474	pisum sativ	199	11	4.0	278	10	09FR13	09FR13	arabidopsis
127	13	4.7	224	10	094C22	094C22	crucifera	200	11	4.0	283	10	094FU0	094FU0	arabidopsis
128	13	4.7	226	10	022584	022584	arabidopsis	201	11	4.0	286	10	09C9G7	09C9G7	arabidopsis
129	13	4.7	226	10	09L8K95	09L8K95	arabidopsis	202	11	4.0	341	10	09FTF4	09FTF4	glycine max
130	13	4.7	243	10	094CJ3	094CJ3	arabidopsis	203	10	3.6	42	10	09X1U1	09X1U1	zea mays (m
131	13	4.7	256	10	094IG1	094IG1	arabis gemm	204	10	3.6	42	10	09S0Q7	09S0Q7	zea mays (m
132	13	4.7	264	10	08S3Y3	08S3Y3	oryza sativ	205	10	3.6	43	10	09S0C5	09S0C5	zea mays (m
133	13	4.7	264	10	09L8C70	09L8C70	oryza sativ	206	10	3.6	43	10	09S0P2	09S0P2	zea mays (m
134	13	4.7	268	10	041843	041843	zea mays (m	207	10	3.6	43	10	09S0B4	09S0B4	zea mays (m
135	13	4.7	271	10	09S8G3	09S8G3	arabidopsis	208	10	3.6	43	10	09S0E5	09S0E5	zea mays (m
136	13	4.7	273	10	093V46	093V46	arabidopsis	209	10	3.6	45	10	08L6S3	08L6S3	gossypium h
137	13	4.7	273	10	094IF8	094IF8	arabidopsis	210	10	3.6	45	10	09S8M7	09S8M7	arabidopsis
138	13	4.7	273	10	094IF6	094IF6	arabidopsis	211	10	3.6	45	10	08L6Q0	08L6Q0	gossypium h
139	13	4.7	273	10	094IF7	094IF7	arabidopsis	212	10	3.6	45	10	049760	049760	arabidopsis
140	13	4.7	273	10	039028	039028	arabidopsis	213	10	3.6	46	10	08L6O5	08L6O5	gossypium h
141	13	4.7	273	10	094IF9	094IF9	arabidopsis	214	10	3.6	46	10	08L6S2	08L6S2	gossypium h
142	13	4.7	273	10	093V20	093V20	arabidopsis	215	10	3.6	47	10	08L6S1	08L6S1	gossypium h
143	13	4.7	285	10	09AR18	09AR18	oryza sativ	216	10	3.6	229	10	049806	049806	arabidopsis
144	13	4.7	293	10	093V35	093V35	oryza sativ	217	10	3.6	263	10	092TE4	092TE4	arabidopsis
145	13	4.7	296	10	09T066	09T066	arabidopsis	218	10	3.6	286	10	08H257	08H257	gossypioide
146	13	4.7	298	10	09S8J7	09S8J7	arabidopsis	219	10	3.6	287	10	08H261	08H261	gossypium h
147	13	4.7	298	10	08H5F5	08H5F5	oryza sativ	220	10	3.6	287	10	08H260	08H260	gossypium h
148	13	4.7	300	10	08W0U3	08W0U3	sorghum bic	221	10	3.6	287	10	08H259	08H259	gossypium r
149	13	4.7	306	10	08W0D0	08W0D0	oryza sativ	222	10	3.6	287	10	08H258	08H258	gossypium h
150	13	4.7	307	10	09FGV3	09FGV3	arabidopsis	223	10	3.6	293	10	049020	049020	gossypium h
151	13	4.7	309	10	049746	049746	arabidopsis	224	10	3.6	293	10	094JN4	094JN4	gossypium h
152	13	4.7	310	10	09M2I9	09M2I9	arabidopsis	225	10	3.6	293	10	094JN5	094JN5	gossypium h
153	13	4.7	315	10	08W1R8	08W1R8	lycopersico	226	10	3.6	317	10	09FP39	09FP39	oryza sativ
154	13	4.7	318	10	08S026	08S026	oryza sativ	227	10	3.6	342	10	09FU07	09FU07	arabidopsis
155	13	4.7	323	10	09LDE1	09LDE1	arabidopsis	228	10	3.6	371	10	022264	022264	arabidopsis
156	13	4.7	329	10	09FG68	09FG68	arabidopsis	229	10	3.6	43	10	09S0F3	09S0F3	zea mays (m
157	13	4.7	333	10	09ZTP6	09ZTP6	arabidopsis	230	9	3.3	43	10	09S7M4	09S7M4	zea mays (m
158	13	4.7	333	10	09S5S0	09S5S0	craterostig	231	9	3.3	43	10	049785	049785	arabidopsis
159	13	4.7	333	10	09F5L1	09F5L1	arabidopsis	232	9	3.3	45	10	049761	049761	arabidopsis
160	13	4.7	335	10	09S5S2	09S5S2	craterostig	233	9	3.3	45	10	065722	065722	arabidopsis
161	13	4.7	374	10	049538	049538	arabidopsis	234	9	3.3	45	10	049800	049800	arabidopsis
162	13	4.7	374	10	049538	049538	arabidopsis	235	9	3.3	45	10	049800	049800	arabidopsis

236	9	3.3	45	10	049771	049771 arabidopsis	309	7	2.6	95	15	090077	090077 human immun
237	9	3.3	106	10	085323	085323 sorghum bic	310	7	2.6	95	15	090079	090079 human immun
238	9	3.3	162	10	085327	085327 zea mays (m	311	7	2.6	95	15	090080	090080 human immun
239	9	3.3	269	10	09SECA	09SECA pimplinella	312	7	2.6	95	15	090078	090078 human immun
240	9	3.3	320	10	09XHVO	09XHVO arabidopsis	313	7	2.6	99	10	085441	085441 sorghum bic
241	9	3.3	330	10	09FR08	09FR08 zea mays (s	314	7	2.6	111	10	043599	043599 oryza sativ
242	9	3.3	331	10	0853X7	0853X7 oryza sativ	315	7	2.6	114	10	0854X1	0854X1 oryza sativ
243	9	3.3	335	10	0944N2	0944N2 zea mays (m	316	7	2.6	118	10	09FXC3	09FXC3 arabidopsis
244	9	3.3	335	10	09FR09	09FR09 zea mays (m	317	7	2.6	119	10	085438	085438 sorghum bic
245	9	3.3	340	10	09FR07	09FR07 zea mays (s	318	7	2.6	122	10	039260	039260 arabidopsis
246	9	3.3	376	10	024579	024579 zea mays (m	319	7	2.6	125	16	09R266	09R266 deinococcus
247	9	3.3	453	10	049019	049019 gossypium h	320	7	2.6	126	2	09ANU7	09ANU7 bradyrhizob
248	8	2.9	45	10	08L6N7	08L6N7 gossypium h	321	7	2.6	127	10	085446	085446 sorghum bic
249	8	2.9	45	10	049772	049772 arabidopsis	322	7	2.6	128	10	085328	085328 zea mays (m
250	8	2.9	45	10	049773	049773 arabidopsis	323	7	2.6	129	10	09FPD2	09FPD2 arabidopsis
251	8	2.9	46	10	08L6M5	08L6M5 gossypium h	324	7	2.6	130	2	09WMY0	09WMY0 cytophaga j
252	8	2.9	46	10	08L6M4	08L6M4 gossypium h	325	7	2.6	130	10	085417	085417 zea mays (m
253	8	2.9	48	10	08L6P2	08L6P2 gossypium h	326	7	2.6	134	16	08CLJ1	08CLJ1 yerstinia pe
254	8	2.9	192	10	08L6R4	08L6R4 cucumis sat	327	7	2.6	135	17	08ZWP4	08ZWP4 pyrobaculum
255	8	2.9	214	10	049780	049780 arabidopsis	328	7	2.6	136	10	P92986	P92986 arabidopsis
256	8	2.9	220	10	09SBG1	09SBG1 arabidopsis	329	7	2.6	139	3	09HGX4	09HGX4 agaricus bi
257	8	2.9	289	16	08NOY6	08NOY6 corynabace	330	7	2.6	146	10	085324	085324 zea mays (m
258	8	2.9	294	10	08LUNG8	08LUNG8 oryza sativ	331	7	2.6	147	16	09JUZ1	09JUZ1 clostridium
259	8	2.9	295	10	064399	064399 arabidopsis	332	7	2.6	150	10	085325	085325 zea mays (m
260	8	2.9	296	10	09A569	09A569 oryza sativ	333	7	2.6	152	10	085400	085400 zea mays (m
261	8	2.9	432	3	096VPO	096VPO kluyveromyc	334	7	2.6	153	10	049649	049649 arabidopsis
262	8	2.9	434	16	09KEZ6	09KEZ6 bacillus ha	335	7	2.6	157	8	08HH88	08HH88 sarcopetes s
263	8	2.9	889	10	08L1C6	08L1C6 oryza sativ	336	7	2.6	157	10	085416	085416 zea mays (m
264	8	2.9	1501	11	08K4X6	08K4X6 mus musculu	337	7	2.6	159	10	085A75	085A75 oryza sativ
265	8	2.6	35	10	08GR19	08GR19 picea maria	338	7	2.6	159	10	085401	085401 zea mays (m
266	7	2.6	42	10	09XIT4	09XIT4 glycine max	339	7	2.6	160	10	085415	085415 zea mays (m
267	7	2.6	43	10	09S0C6	09S0C6 zea mays (m	340	7	2.6	161	5	08ILK3	08ILK3 plasmodium
268	7	2.6	45	10	08L6M4	08L6M4 gossypium h	341	7	2.6	162	10	0853Y7	0853Y7 sorghum bic
269	7	2.6	45	10	049789	049789 arabidopsis	342	7	2.6	162	15	08BDX4	08BDX4 human immun
270	7	2.6	45	10	08L6N0	08L6N0 gossypium h	343	7	2.6	163	8	03S592	03S592 physearum po
271	7	2.6	45	10	09SAL7	09SAL7 arabidopsis	344	7	2.6	163	10	09LRC3	09LRC3 titilicium ae
272	7	2.6	45	10	049755	049755 arabidopsis	345	7	2.6	166	10	085322	085322 sorghum bic
273	7	2.6	45	10	08L6R2	08L6R2 gossypium h	346	7	2.6	172	16	08PRL6	08PRL6 xanthomonas
274	7	2.6	45	10	049776	049776 arabidopsis	347	7	2.6	173	16	08PRL6	08PRL6 corynabace
275	7	2.6	46	10	08L541	08L541 gossypium h	348	7	2.6	174	2	045501	045501 bacillus su
276	7	2.6	50	10	08L6Q7	08L6Q7 gossypium h	349	7	2.6	174	10	08L438	08L438 oryza sativ
277	7	2.6	61	10	085429	085429 zea mays (m	350	7	2.6	176	10	085A74	085A74 oryza sativ
278	7	2.6	63	10	085436	085436 sorghum bic	351	7	2.6	184	10	0853Y8	0853Y8 sorghum bic
279	7	2.6	63	10	085434	085434 sorghum bic	352	7	2.6	192	15	08UT78	08UT78 human immun
280	7	2.6	63	10	085435	085435 sorghum bic	353	7	2.6	194	5	08SPH1	08SPH1 encephalito
281	7	2.6	64	10	085439	085439 sorghum bic	354	7	2.6	198	10	08RUC5	08RUC5 zea mays (m
282	7	2.6	64	10	085442	085442 sorghum bic	355	7	2.6	199	16	08EVC1	08EVC1 mycoplasma
283	7	2.6	64	10	085437	085437 sorghum bic	356	7	2.6	204	10	0853Y1	0853Y1 oryza sativ
284	7	2.6	65	10	085440	085440 sorghum bic	357	7	2.6	204	10	0853Y1	0853Y1 oryza sativ
285	7	2.6	67	10	043597	043597 oryza sativ	358	7	2.6	204	10	0853Y2	0853Y2 oryza sativ
286	7	2.6	71	10	08H1Z4	08H1Z4 lycopersico	359	7	2.6	208	10	085402	085402 zea mays (m
287	7	2.6	71	10	08H1Z3	08H1Z3 lycopersico	360	7	2.6	209	16	09CK25	09CK25 pasteurella
288	7	2.6	83	10	08LKG4	08LKG4 lolium pere	361	7	2.6	209	16	08ZCF0	08ZCF0 yerstinia pe
289	7	2.6	87	10	085424	085424 zea mays (m	362	7	2.6	210	10	085421	085421 zea mays (m
290	7	2.6	87	10	085443	085443 sorghum bic	363	7	2.6	212	10	085329	085329 zea mays (m
291	7	2.6	87	10	085431	085431 zea mays (m	364	7	2.6	212	10	09S0S8	09S0S8 arabidopsis
292	7	2.6	87	10	085445	085445 sorghum bic	365	7	2.6	212	10	0853Y4	0853Y4 oryza sativ
293	7	2.6	87	10	085419	085419 zea mays (m	366	7	2.6	215	10	0853X9	0853X9 oryza sativ
294	7	2.6	87	10	085432	085432 zea mays (m	367	7	2.6	215	16	08ZNI1	08ZNI1 salmoneila
295	7	2.6	87	10	085428	085428 zea mays (m	368	7	2.6	215	16	085766	085766 salmoneila
296	7	2.6	87	10	085433	085433 sorghum bic	369	7	2.6	215	16	08XEB6	08XEB6 escherichia
297	7	2.6	87	10	085426	085426 zea mays (m	370	7	2.6	216	10	09LUB4	09LUB4 arabidopsis
298	7	2.6	88	10	085423	085423 zea mays (m	371	7	2.6	218	16	09JOT9	09JOT9 neisseria m
299	7	2.6	88	10	085430	085430 zea mays (m	372	7	2.6	220	10	09PEV3	09PEV3 oryza sativ
300	7	2.6	92	10	09JMB4	09JMB4 arabidopsis	373	7	2.6	220	16	08D0X7	08D0X7 yerstinia pe
301	7	2.6	95	15	090074	090074 human immun	374	7	2.6	221	10	094CT3	094CT3 oryza sativ
302	7	2.6	95	15	090072	090072 human immun	375	7	2.6	223	10	085G47	085G47 oryza sativ
303	7	2.6	95	15	090082	090082 human immun	376	7	2.6	224	17	08RT01	08RT01 methanopyru
304	7	2.6	95	15	090081	090081 human immun	377	7	2.6	225	16	08BFR5	08BFR5 escherichia
305	7	2.6	95	15	090071	090071 human immun	378	7	2.6	228	5	08SUT1	08SUT1 encephalito
306	7	2.6	95	15	090073	090073 human immun	379	7	2.6	229	10	0853Y0	0853Y0 oryza sativ
307	7	2.6	95	15	090076	090076 human immun	380	7	2.6	233	10	09FYX6	09FYX6 arabidopsis
308	7	2.6	95	15	090075	090075 human immun	381	7	2.6	233	16	08RDY1	08RDY1 fueobacteri

382	7	2.6	234	10	050069	050069 arabidopsis	455	7	2.6	286	10	036PG1	036PG1 arabidopsis
383	7	2.6	235	10	09FN86	09FN86 arabidopsis	456	7	2.6	287	10	08FR09	08FR09 oryza sativ
384	7	2.6	238	10	09SCP1	09SCP1 arabidopsis	457	7	2.6	287	10	08FR08	08FR08 arabidopsis
385	7	2.6	239	10	09LTV4	09LTV4 arabidopsis	458	7	2.6	288	10	09CA63	09CA63 hordeum vul
386	7	2.6	239	10	023889	023889 oryza sativ	459	7	2.6	296	10	09SGU3	09SGU3 arabidopsis
387	7	2.6	239	16	08FRG8	08FRG8 corynebacte	460	7	2.6	297	10	09M213	09M213 arabidopsis
388	7	2.6	242	10	09FELS	09FELS nicotiana t	461	7	2.6	299	10	09M2D9	09M2D9 arabidopsis
389	7	2.6	243	10	09FX36	09FX36 arabidopsis	462	7	2.6	300	16	08EGK6	08EGK6 shewaniella
390	7	2.6	243	10	09ZTD7	09ZTD7 arabidopsis	463	7	2.6	301	2	09KZV8	09KZV8 campylobact
391	7	2.6	244	10	09ZTD7	09ZTD7 arabidopsis	464	7	2.6	301	2	09KZV8	09KZV8 campylobact
392	7	2.6	245	10	09FN72	09FN72 oryza sativ	465	7	2.6	301	2	09KZV8	09KZV8 campylobact
393	7	2.6	245	10	09SNW9	09SNW9 arabidopsis	466	7	2.6	301	17	09GN12	09GN12 arabidopsis
394	7	2.6	246	10	039153	039153 arabidopsis	467	7	2.6	301	17	09GN12	09GN12 arabidopsis
395	7	2.6	246	10	09LNC9	09LNC9 arabidopsis	468	7	2.6	302	10	09LNR7	09LNR7 oryza sativ
396	7	2.6	248	5	097008	097008 leishmania	469	7	2.6	304	10	039155	039155 arabidopsis
397	7	2.6	248	10	09LMT5	09LMT5 arabidopsis	470	7	2.6	304	10	049745	049745 arabidopsis
398	7	2.6	249	10	09ZTD9	09ZTD9 arabidopsis	471	7	2.6	305	10	09FDW1	09FDW1 arabidopsis
399	7	2.6	249	10	08LCC8	08LCC8 arabidopsis	472	7	2.6	305	10	09A1B3	09A1B3 arabidopsis
400	7	2.6	249	10	09SJT8	09SJT8 arabidopsis	473	7	2.6	305	10	08LBC5	08LBC5 arabidopsis
401	7	2.6	249	16	08XN17	08XN17 clostridium	474	7	2.6	305	10	039154	039154 arabidopsis
402	7	2.6	250	10	09STM6	09STM6 arabidopsis	475	7	2.6	306	10	09FZ14	09FZ14 solanum tub
403	7	2.6	250	10	09SZC2	09SZC2 arabidopsis	476	7	2.6	307	10	09ASFS	09ASFS oryza sativ
404	7	2.6	251	10	006944	006944 hordeum vul	477	7	2.6	307	10	08W389	08W389 oryza sativ
405	7	2.6	253	10	08LSN8	08LSN8 vitis labru	478	7	2.6	307	10	09LTV2	09LTV2 arabidopsis
406	7	2.6	253	10	08LSN7	08LSN7 vitis labru	479	7	2.6	307	10	08H0H3	08H0H3 nicotiana t
407	7	2.6	255	10	09FZ13	09FZ13 solanum tub	480	7	2.6	309	10	08LA62	08LA62 arabidopsis
408	7	2.6	255	10	09FZ13	09FZ13 arabidopsis	481	7	2.6	309	10	022179	022179 arabidopsis
409	7	2.6	256	10	08H6E9	08H6E9 sauburea m	482	7	2.6	309	10	09ATD4	09ATD4 gossypium h
410	7	2.6	257	10	004140	004140 oryza sativ	483	7	2.6	309	10	08H0P2	08H0P2 oryza sativ
411	7	2.6	260	10	09A667	09A667 oryza sativ	484	7	2.6	310	10	09FPJ2	09FPJ2 arabidopsis
412	7	2.6	261	10	09SMT1	09SMT1 arabidopsis	485	7	2.6	311	10	004109	004109 oryza sativ
413	7	2.6	261	10	09T0G9	09T0G9 arabidopsis	486	7	2.6	314	10	09CC95	09CC95 arabidopsis
414	7	2.6	262	17	08TUK3	08TUK3 methanosarc	487	7	2.6	316	10	08I392	08I392 anticirrhium
415	7	2.6	262	17	08PX61	08PX61 methanosarc	488	7	2.6	317	10	09LS17	09LS17 arabidopsis
416	7	2.6	263	10	09LY95	09LY95 arabidopsis	489	7	2.6	319	10	09SE66	09SE66 arabidopsis
417	7	2.6	264	10	09S7E3	09S7E3 glycine max	490	7	2.6	320	10	023160	023160 arabidopsis
418	7	2.6	264	10	09XIU9	09XIU9 glycine max	491	7	2.6	321	10	09LDD5	09LDD5 arabidopsis
419	7	2.6	265	10	040174	040174 lycopersico	492	7	2.6	321	10	038739	038739 anticirrhium
420	7	2.6	266	10	094GN6	094GN6 arabidopsis	493	7	2.6	323	10	09S8G3	09S8G3 arabidopsis
421	7	2.6	267	10	09LTM9	09LTM9 oryza sativ	494	7	2.6	323	10	09SCU7	09SCU7 arabidopsis
422	7	2.6	268	10	09LTR5	09LTR5 arabidopsis	495	7	2.6	323	10	08LBP8	08LBP8 arabidopsis
423	7	2.6	268	10	08I394	08I394 anticirrhium	496	7	2.6	324	10	09M0Y5	09M0Y5 arabidopsis
424	7	2.6	269	10	09LKI4	09LKI4 arabidopsis	497	7	2.6	325	10	09LPE1	09LPE1 arabidopsis
425	7	2.6	269	10	042379	042379 arabidopsis	498	7	2.6	326	10	09LXFL	09LXFL arabidopsis
426	7	2.6	272	10	09XIU5	09XIU5 glycine max	499	7	2.6	327	5	09GPO4	09GPO4 trypanosoma
427	7	2.6	272	10	08LEE5	08LEE5 arabidopsis	500	7	2.6	327	10	064534	064534 arabidopsis
428	7	2.6	273	10	049744	049744 arabidopsis	501	7	2.6	327	10	09SE67	09SE67 arabidopsis
429	7	2.6	274	10	09SA47	09SA47 arabidopsis	502	7	2.6	329	10	08H0H0	08H0H0 nicotiana t
430	7	2.6	274	10	049608	049608 arabidopsis	503	7	2.6	330	10	09CA52	09CA52 arabidopsis
431	7	2.6	274	10	040173	040173 lycopersico	504	7	2.6	330	10	09SE24	09SE24 arabidopsis
432	7	2.6	274	10	09SM28	09SM28 arabidopsis	505	7	2.6	331	10	043524	043524 lycopersico
433	7	2.6	274	10	08I395	08I395 anticirrhium	506	7	2.6	332	10	042467	042467 arabidopsis
434	7	2.6	275	10	09XIU8	09XIU8 glycine max	507	7	2.6	333	10	09SN78	09SN78 arabidopsis
435	7	2.6	276	10	09LUR5	09LUR5 arabidopsis	508	7	2.6	333	10	09C9C8	09C9C8 arabidopsis
436	7	2.6	276	10	004108	004108 oryza sativ	509	7	2.6	334	10	09S8F3	09S8F3 arabidopsis
437	7	2.6	277	10	09SLT1	09SLT1 nicotiana t	510	7	2.6	335	10	09ZTC1	09ZTC1 arabidopsis
438	7	2.6	278	10	09J391	09J391 nicotiana t	511	7	2.6	336	10	09FLR1	09FLR1 arabidopsis
439	7	2.6	278	10	049759	049759 arabidopsis	512	7	2.6	336	10	09M118	09M118 arabidopsis
440	7	2.6	279	10	004110	004110 oryza sativ	513	7	2.6	336	10	049765	049765 arabidopsis
441	7	2.6	280	10	09CS59	09CS59 arabidopsis	514	7	2.6	337	10	09M8G3	09M8G3 arabidopsis
442	7	2.6	280	10	09SBG0	09SBG0 arabidopsis	515	7	2.6	338	10	09S7K6	09S7K6 oryza sativ
443	7	2.6	280	10	002993	002993 petunia hyb	516	7	2.6	338	10	09SPG5	09SPG5 arabidopsis
444	7	2.6	280	10	09LXV2	09LXV2 arabidopsis	517	7	2.6	341	10	08GRH7	08GRH7 oryza sativ
445	7	2.6	280	10	08GYP5	08GYP5 arabidopsis	518	7	2.6	343	10	09ZRY6	09ZRY6 arabidopsis
446	7	2.6	281	10	09SLT0	09SLT0 nicotiana t	519	7	2.6	343	10	09CBU1	09CBU1 arabidopsis
447	7	2.6	281	10	09SLT2	09SLT2 nicotiana t	520	7	2.6	343	10	09LZK4	09LZK4 arabidopsis
448	7	2.6	282	5	094011	094011 caenorhabdi	521	7	2.6	345	10	09LE63	09LE63 lycopersico
449	7	2.6	282	10	09C7U7	09C7U7 arabidopsis	522	7	2.6	347	10	040175	040175 lycopersico
450	7	2.6	282	10	09M0U5	09M0U5 arabidopsis	523	7	2.6	348	10	09ZS14	09ZS14 arabidopsis
451	7	2.6	282	10	09MOK4	09MOK4 arabidopsis	524	7	2.6	349	10	09SPG8	09SPG8 arabidopsis
452	7	2.6	283	10	09SUI1	09SUI1 arabidopsis	525	7	2.6	350	10	09SKB6	09SKB6 arabidopsis
453	7	2.6	285	10	09LTC4	09LTC4 arabidopsis	526	7	2.6	350	10	08W4E0	08W4E0 arabidopsis
454	7	2.6	285	10	08GRY3	08GRY3 oryza sativ	527	7	2.6	350	10	08LBF0	08LBF0 arabidopsis



528	7	2.6	350	10	Q9LDR8	Q9LDR8 arabidopsis	601	7	2.6	566	11	Q9ERQ3	Q9ERQ3 mus musculus
529	7	2.6	351	10	Q9SPG0	Q9SPG0 arabidopsis	602	7	2.6	581	10	Q9FIA4	Q9FIA4 arabidopsis
530	7	2.6	351	10	Q9ZMZ5	Q9ZMZ5 glycine max	603	7	2.6	584	13	Q9PDK5	Q9PDK5 carassius a
531	7	2.6	352	10	Q9LOE2	Q9LOE2 arabidopsis	604	7	2.6	587	10	Q9ZWP9	Q9ZWP9 cucumis mel
532	7	2.6	352	10	Q92685	Q92685 arabidopsis	605	7	2.6	588	4	Q8WYC3	Q8WYC3 homo sapien
533	7	2.6	352	10	Q8LPH6	Q8LPH6 arabidopsis	606	7	2.6	588	10	Q947F0	Q947F0 malus domes
534	7	2.6	352	10	Q65409	Q65409 arabidopsis	607	7	2.6	590	10	Q8VZ70	Q8VZ70 arabidopsis
535	7	2.6	352	10	Q49782	Q49782 arabidopsis	608	7	2.6	608	10	Q93X77	Q93X77 malus domes
536	7	2.6	352	10	Q9ZNZ6	Q9ZNZ6 glycine max	609	7	2.6	608	10	Q93W76	Q93W76 malus domes
537	7	2.6	357	10	Q8RUN0	Q8RUN0 oryza sativ	610	7	2.6	639	5	Q9VXX5	Q9VXX5 drosophila
538	7	2.6	357	10	Q9ZUM6	Q9ZUM6 arabidopsis	611	7	2.6	646	10	Q8S3R9	Q8S3R9 oryza sativ
539	7	2.6	357	10	Q8G0V1	Q8G0V1 oryza sativ	612	7	2.6	648	3	Q13526	Q13526 saccharomyc
540	7	2.6	358	10	Q9LV10	Q9LV10 arabidopsis	613	7	2.6	669	17	Q58474	Q58474 pyrococcus
541	7	2.6	359	10	Q9LOX5	Q9LOX5 arabidopsis	614	7	2.6	670	10	Q9LV04	Q9LV04 arabidopsis
542	7	2.6	359	10	Q9XFM9	Q9XFM9 anticirrhium	615	7	2.6	697	3	Q74504	Q74504 schizosach
543	7	2.6	360	10	Q8GMP0	Q8GMP0 arabidopsis	616	7	2.6	697	5	Q81J52	Q81J52 plasmodium
544	7	2.6	364	10	Q9FZ15	Q9FZ15 solanum tub	617	7	2.6	702	17	Q26412	Q26412 methanobact
545	7	2.6	365	10	Q9S9Z2	Q9S9Z2 arabidopsis	618	7	2.6	714	16	Q8YRK3	Q8YRK3 anabaena sp
546	7	2.6	366	10	Q9SPG2	Q9SPG2 arabidopsis	619	7	2.6	745	2	Q9WM95	Q9WM95 salmonella
547	7	2.6	366	10	Q8VZQ2	Q8VZQ2 arabidopsis	620	7	2.6	745	2	Q9Z4C1	Q9Z4C1 shigella so
548	7	2.6	367	10	Q9LHY6	Q9LHY6 oryza sativ	621	7	2.6	745	10	Q23618	Q23618 arabidopsis
549	7	2.6	367	10	Q04192	Q04192 oryza sativ	622	7	2.6	767	10	Q8S8H6	Q8S8H6 arabidopsis
550	7	2.6	367	10	Q9SPG3	Q9SPG3 arabidopsis	623	7	2.6	822	10	Q41370	Q41370 spinacta ol
551	7	2.6	368	10	Q04141	Q04141 oryza sativ	624	7	2.6	833	5	Q44476	Q44476 caenorhabdi
552	7	2.6	369	10	Q65249	Q65249 arabidopsis	625	7	2.6	840	15	Q8QDM5	Q8QDM5 human immun
553	7	2.6	370	10	Q9SRB0	Q9SRB0 arabidopsis	626	7	2.6	848	15	Q8UTN5	Q8UTN5 human immun
554	7	2.6	370	10	Q8EBH3	Q8EBH3 shewanella	627	7	2.6	866	10	Q9M9E2	Q9M9E2 arabidopsis
555	7	2.6	375	10	Q8H220	Q8H220 populus x c	628	7	2.6	901	5	Q817T4	Q817T4 trypanosoma
556	7	2.6	376	17	Q8PVP6	Q8PVP6 methanosc	629	7	2.6	938	16	Q67811	Q67811 aquifex aeo
557	7	2.6	386	16	Q97K26	Q97K26 clostridium	630	7	2.6	954	5	Q9V077	Q9V077 drosophila
558	7	2.6	389	10	Q9S773	Q9S773 arabidopsis	631	7	2.6	956	16	Q84812	Q84812 chlamydia t
559	7	2.6	393	10	Q42575	Q42575 arabidopsis	632	7	2.6	1000	3	Q9Y802	Q9Y802 schizosach
560	7	2.6	395	10	Q23893	Q23893 oryza sativ	633	7	2.6	1082	5	Q81TC9	Q81TC9 plasmodium
561	7	2.6	399	10	Q9LDM5	Q9LDM5 arabidopsis	634	7	2.6	1082	5	Q81J86	Q81J86 plasmodium
562	7	2.6	400	10	Q8LAR8	Q8LAR8 oryza sativ	635	7	2.6	1083	3	Q94189	Q94189 cryptococcu
563	7	2.6	403	16	Q8DBB9	Q8DBB9 vibrio vuln	636	7	2.6	1156	16	Q66878	Q66878 aquifex aeo
564	7	2.6	412	5	Q8SVX8	Q8SVX8 encephalit	637	7	2.6	1175	10	Q8S566	Q8S566 oryza sativ
565	7	2.6	413	10	Q8S282	Q8S282 oryza sativ	638	7	2.6	1200	10	Q9LCZ7	Q9LCZ7 arabidopsis
566	7	2.6	414	10	Q9SAV9	Q9SAV9 arabidopsis	639	7	2.6	1204	10	Q9S9K0	Q9S9K0 arabidopsis
567	7	2.6	417	16	Q9CUX0	Q9CUX0 paeteurella	640	7	2.6	1285	5	Q96195	Q96195 plasmodium
568	7	2.6	418	10	Q9ZWC3	Q9ZWC3 arabidopsis	641	7	2.6	1297	10	Q9XEM9	Q9XEM9 oryza sativ
569	7	2.6	421	10	Q02992	Q02992 petunia hyb	642	7	2.6	1337	5	Q9Y0E8	Q9Y0E8 plasmodium
570	7	2.6	427	10	Q9S1M1	Q9S1M1 arabidopsis	643	7	2.6	1397	5	Q81BA3	Q81BA3 plasmodium
571	7	2.6	428	10	Q9A0V2	Q9A0V2 oryza sativ	644	7	2.6	1620	10	Q8H801	Q8H801 oryza sativ
572	7	2.6	443	2	Q30867	Q30867 haemophilus	645	7	2.6	1710	5	Q81239	Q81239 plasmodium
573	7	2.6	445	5	Q9U8U8	Q9U8U8 brachiocto	646	7	2.6	1771	12	Q93357	Q93357 banana mild
574	7	2.6	450	17	Q8TCY3	Q8TCY3 methanopyru	647	7	2.6	1788	12	Q8UVV6	Q8UVV6 oryza sativ
575	7	2.6	451	10	Q9LHS6	Q9LHS6 arabidopsis	648	7	2.6	1824	5	Q81HV9	Q81HV9 pseudomonas
576	7	2.6	470	10	Q9FR42	Q9FR42 nicotiana t	649	7	2.6	1883	5	Q24376	Q24376 drosophila
577	7	2.6	474	10	Q9FR43	Q9FR43 nicotiana t	650	7	2.6	2136	10	Q8RYW8	Q8RYW8 oryza sativ
578	7	2.6	480	10	Q8W1M5	Q8W1M5 arabidopsis	651	7	2.6	2254	12	Q98EJ5	Q98EJ5 porcine ent
579	7	2.6	490	10	Q80883	Q80883 arabidopsis	652	7	2.6	2352	16	Q9H1R8	Q9H1R8 pseudomonas
580	7	2.6	493	12	Q83449	Q83449 murine roca	653	7	2.6	2360	3	Q94188	Q94188 cryptococcu
581	7	2.6	514	5	Q21458	Q21458 caenorhabdi	654	7	2.6	2755	5	Q81HP8	Q81HP8 plasmodium
582	7	2.6	517	10	Q02994	Q02994 petunia hyb	655	7	2.6	2766	8	Q8GM85	Q8GM85 symbiont ba
583	7	2.6	520	10	Q8W1W6	Q8W1W6 arabidopsis	656	7	2.6	278	8	Q8W1J4	Q8W1J4 picea maria
584	7	2.6	523	10	Q94FL7	Q94FL7 arabidopsis	657	7	2.6	35	10	Q8GUX4	Q8GUX4 picea maria
585	7	2.6	525	11	Q99PS6	Q99PS6 mus musculu	658	7	2.6	35	10	Q8GUM9	Q8GUM9 picea maria
586	7	2.6	525	11	Q99PS6	Q99PS6 mus musculu	659	7	2.6	35	10	Q8GUM7	Q8GUM7 picea maria
587	7	2.6	525	11	Q9ESB3	Q9ESB3 mus musculu	660	7	2.6	38	4	Q96KF8	Q96KF8 homo sapien
588	7	2.6	527	5	Q9XVNB	Q9XVNB caenorhabdi	661	7	2.6	40	12	Q9JZ27	Q9JZ27 hepatitis c
589	7	2.6	531	10	Q49601	Q49601 arabidopsis	662	7	2.6	42	10	Q9S085	Q9S085 arabidopsis
590	7	2.6	531	12	Q09642	Q09642 influenza	663	7	2.6	42	10	Q9S083	Q9S083 arabidopsis
591	7	2.6	535	15	Q8QDX5	Q8QDX5 human immun	664	7	2.6	42	10	Q9X1T3	Q9X1T3 glycine max
592	7	2.6	546	10	Q9XFN2	Q9XFN2 avena sativ	665	7	2.6	42	10	Q9X1T9	Q9X1T9 glycine max
593	7	2.6	548	10	Q9XEP0	Q9XEP0 lolium tenu	666	7	2.6	42	10	Q9X1U2	Q9X1U2 glycine max
594	7	2.6	550	15	Q8ODX3	Q8ODX3 human immun	667	7	2.6	42	10	Q9SXF9	Q9SXF9 glycine max
595	7	2.6	553	10	P93417	P93417 oryza sativ	668	7	2.6	42	10	Q9S084	Q9S084 arabidopsis
596	7	2.6	553	10	Q9FR97	Q9FR97 arabidopsis	669	7	2.6	43	10	Q9S095	Q9S095 zea mays (m
597	7	2.6	553	10	Q8SA38	Q8SA38 oryza sativ	670	7	2.6	43	10	Q9S0E8	Q9S0E8 zea mays (m
598	7	2.6	553	10	Q9ARM1	Q9ARM1 hordeum vul	671	7	2.6	43	10	Q9S0B0	Q9S0B0 zea mays (m
599	7	2.6	553	10	Q96464	Q96464 hordeum vul	672	7	2.6	43	10	Q9S0B1	Q9S0B1 zea mays (m
600	7	2.6	562	5	Q813P8	Q813P8 plasmodium	673	7	2.6	43	10	Q9S0B6	Q9S0B6 zea mays (m

674	6	2.2	43	10	Q9SQF0	Q9sqf0 zea mays (m	747	6	2.2	83	8	Q36875	Q36875 hylobates c
675	6	2.2	43	10	Q9SQB8	Q9sqb8 zea mays (m	748	6	2.2	83	15	Q87627	Q87627 chimpanzee
676	6	2.2	43	10	Q9SQC2	Q9sqc2 zea mays (m	749	6	2.2	87	5	Q21168	Q21168 caenorhabdi
677	6	2.2	43	10	Q9SQC1	Q9sqc1 zea mays (m	750	6	2.2	87	10	Q85005	Q85005 oryza sativ
678	6	2.2	43	10	Q9SQD3	Q9sqd3 zea mays (m	751	6	2.2	87	12	Q41126	Q41126 paramecium
679	6	2.2	43	10	Q9SQD0	Q9sqd0 zea mays (m	752	6	2.2	88	2	Q9K1M0	Q9K1M0 streptomyce
680	6	2.2	43	10	Q9SQC0	Q9sqc0 zea mays (m	753	6	2.2	88	4	Q94795	Q94795 homo sapien
681	6	2.2	43	10	Q9SQC4	Q9sqc4 zea mays (m	754	6	2.2	88	15	Q70146	Q70146 human immun
682	6	2.2	43	10	Q9SQE7	Q9sqe7 zea mays (m	755	6	2.2	89	17	Q8P253	Q8P253 mechanosarc
683	6	2.2	43	10	Q9SQD7	Q9sqd7 zea mays (m	756	6	2.2	90	17	Q8PRR8	Q8PRR8 mechanosarc
684	6	2.2	43	10	Q9SQ91	Q9sq91 zea mays (m	757	6	2.2	91	15	Q78772	Q78772 human immun
685	6	2.2	43	10	Q9SQ96	Q9sq96 zea mays (m	758	6	2.2	91	15	Q9Q0C9	Q9Q0C9 human immun
686	6	2.2	43	10	Q9SQ95	Q9sq95 zea mays (m	759	6	2.2	92	6	P79309	P79309 sus scrofa
687	6	2.2	43	10	Q9SQB5	Q9sqb5 zea mays (m	760	6	2.2	93	15	Q69883	Q69883 human immun
688	6	2.2	43	10	Q9SQB9	Q9sqb9 zea mays (m	761	6	2.2	94	8	Q9MTM1	Q9MTM1 oenothera h
689	6	2.2	43	10	Q9SQD6	Q9sqd6 zea mays (m	762	6	2.2	95	12	Q83003	Q83003 lymphocytic
690	6	2.2	43	10	Q9S7Y8	Q9s7y8 zea mays (m	763	6	2.2	95	17	Q8U112	Q8U112 pyrococcus
691	6	2.2	43	10	Q9S7F1	Q9s7f1 zea mays (m	764	6	2.2	96	17	Q8TYM3	Q8TYM3 methanopyru
692	6	2.2	43	10	Q9SQE6	Q9sqe6 zea mays (m	765	6	2.2	98	8	Q9MTA1	Q9MTA1 cythionecte
693	6	2.2	43	10	Q9SQ99	Q9sq99 zea mays (m	766	6	2.2	100	9	Q94M08	Q94M08 bacterioph
694	6	2.2	43	10	Q9S704	Q9s704 zea mays (m	767	6	2.2	101	16	Q8Y9G0	Q8Y9G0 bruceella me
695	6	2.2	43	10	Q9SQD4	Q9sqd4 zea mays (m	768	6	2.2	101	16	Q9F8P0	Q9F8P0 streptomyce
696	6	2.2	43	10	Q9SQ89	Q9sq89 zea mays (m	769	6	2.2	103	12	Q9E8F7	Q9E8F7 porcine ade
697	6	2.2	43	10	Q9SQA1	Q9sqa1 zea mays (m	770	6	2.2	103	16	Q9RTT9	Q9RTT9 deinococcus
698	6	2.2	43	10	Q9SQ97	Q9sq97 zea mays (m	771	6	2.2	105	11	Q9CRG3	Q9CRG3 mus musculu
699	6	2.2	43	10	Q9SQA7	Q9sqa7 zea mays (m	772	6	2.2	105	11	Q9D256	Q9D256 mus musculu
700	6	2.2	43	10	Q9SQD2	Q9sqd2 zea mays (m	773	6	2.2	107	10	P93276	P93276 arabidopsis
701	6	2.2	43	10	Q9SQA6	Q9sqa6 zea mays (m	774	6	2.2	108	2	Q938A7	Q938A7 staphylococ
702	6	2.2	43	10	Q9SQA0	Q9sqa0 zea mays (m	775	6	2.2	108	5	P91957	P91957 helix lincor
703	6	2.2	43	10	Q9SQD8	Q9sqd8 zea mays (m	776	6	2.2	108	13	Q8UMG0	Q8UMG0 salmo salar
704	6	2.2	43	10	Q9SQA4	Q9sqa4 zea mays (m	777	6	2.2	109	16	Q9CK33	Q9CK33 pasteurella
705	6	2.2	43	10	Q9SQB3	Q9sqb3 zea mays (m	778	6	2.2	109	16	Q9B892	Q9B892 rhizobium l
706	6	2.2	43	10	Q9SQB2	Q9sqb2 zea mays (m	779	6	2.2	112	16	Q915R5	Q915R5 pseudomonas
707	6	2.2	43	10	Q9SQD1	Q9sqd1 zea mays (m	780	6	2.2	113	17	Q9Y9X8	Q9Y9X8 aeropyrum p
708	6	2.2	43	10	Q9SQA5	Q9sqa5 zea mays (m	781	6	2.2	114	17	Q8RT61	Q8RT61 methanosarc
709	6	2.2	45	8	Q32928	Q32928 pinus thunb	782	6	2.2	114	17	Q8PUS2	Q8PUS2 methanosarc
710	6	2.2	45	10	Q8L6L9	Q8l6l9 gossypium h	783	6	2.2	115	16	Q8DU03	Q8DU03 streptococ
711	6	2.2	45	10	Q9SAL6	Q9sala6 arabidopsis	784	6	2.2	116	6	Q9MYT2	Q9MYT2 oryctolagus
712	6	2.2	45	10	Q49754	Q49754 arabidopsis	785	6	2.2	117	12	Q8V7D2	Q8v7d2 ct virus . o
713	6	2.2	45	10	Q49777	Q49777 arabidopsis	786	6	2.2	117	12	Q8UYC8	Q8uyc8 bt virus . o
714	6	2.2	45	10	Q49805	Q49805 arabidopsis	787	6	2.2	117	16	Q8Z7T5	Q8z7t5 salmoneila
715	6	2.2	45	10	Q49781	Q49781 arabidopsis	788	6	2.2	118	10	Q9ZTDS	Q9ztds arabidopsis
716	6	2.2	45	10	Q8L6O6	Q8l6o6 gossypium h	789	6	2.2	119	15	Q71347	Q71347 human immun
717	6	2.2	45	10	Q49802	Q49802 arabidopsis	790	6	2.2	119	17	Q9YE26	Q9YE26 aeropyrum p
718	6	2.2	45	10	Q9SAL9	Q9sala9 arabidopsis	791	6	2.2	121	16	Q9ZBJ6	Q9ZBJ6 listeria in
719	6	2.2	45	10	Q9SAM6	Q9sam6 arabidopsis	792	6	2.2	121	16	Q9RO11	Q9roj1 listeria mo
720	6	2.2	45	10	Q49779	Q49779 arabidopsis	793	6	2.2	121	16	Q8UYF3	Q8uyf3 agrobacteri
721	6	2.2	45	10	Q8L6R4	Q8l6r4 arabidopsis	794	6	2.2	123	13	Q98AU8	Q98au8 rhizobium l
722	6	2.2	46	10	Q8L6R4	Q8l6r4 gossypium h	795	6	2.2	125	5	Q45212	Q45212 brugia paha
723	6	2.2	46	10	Q8L6L8	Q8l6l8 gossypium h	796	6	2.2	125	10	Q8BHU4	Q8bhu4 mus musculu
724	6	2.2	46	10	Q8L6R0	Q8l6r0 gossypium h	797	6	2.2	127	11	Q8BHU2	Q8bhu2 mus musculu
725	6	2.2	48	16	Q8NKK7	Q8nkk7 staphylococ	798	6	2.2	128	5	Q9NA43	Q9na43 caenorhabdi
726	6	2.2	53	12	Q9IER1	Q9ier1 citrus tris	799	6	2.2	128	5	Q8SVP5	Q8svp5 encephalito
727	6	2.2	55	10	Q9SE06	Q9seo6 glycine max	800	6	2.2	128	10	Q9SMN5	Q9smn5 dubautia pl
728	6	2.2	56	10	Q94FW4	Q94fw4 sorghum bic	801	6	2.2	128	10	Q9SMN3	Q9smn3 dubautia ra
729	6	2.2	56	10	Q8LKH8	Q8lkh8 lolium pere	802	6	2.2	128	10	Q9SWM3	Q9swm3 argyroxiphi
730	6	2.2	55	16	Q8DP94	Q8dp94 vibrio vuln	803	6	2.2	128	10	Q9SMM2	Q9smm2 wilkesia gy
731	6	2.2	57	10	Q38974	Q38974 arabidopsis	804	6	2.2	128	10	Q9SWM1	Q9swm1 dubautia sc
732	6	2.2	57	16	Q8NPF7	Q8npf7 corynebacte	805	6	2.2	128	10	Q9SMN7	Q9smn7 dubautia la
733	6	2.2	60	8	Q9TA46	Q9ta46 elephas max	806	6	2.2	128	10	Q9SMP1	Q9smp1 argyroxiphi
734	6	2.2	61	16	Q8G2W7	Q8g2w7 bruceella su	807	6	2.2	128	10	Q9SMM4	Q9smm4 antilocaprus
735	6	2.2	61	17	Q8PUX1	Q8pux1 methanosarc	808	6	2.2	128	15	Q8AON7	Q8aon7 human immun
736	6	2.2	62	16	Q8FTAI	Q8ftai escherichia	809	6	2.2	128	16	Q8AGD8	Q8agd8 bruceella me
737	6	2.2	64	16	Q8X2V0	Q8x2v0 escherichia	810	6	2.2	129	2	Q9L6L0	Q9l6l0 vibrio fisc
738	6	2.2	69	17	Q8TLR1	Q8tlr1 methanosarc	811	6	2.2	130	2	Q9L7B6	Q9l7b6 borrelia bu
739	6	2.2	70	4	Q9HD35	Q9hd35 homo sapien	812	6	2.2	131	10	Q96475	Q96475 lupinus ang
740	6	2.2	71	16	Q9CJ09	Q9cjd9 pasteurella	813	6	2.2	132	1	Q8XFE4	Q8xfz4 salmoneila
741	6	2.2	71	12	Q8O070	Q8o070 camelopox vi	814	6	2.2	132	2	Q9RNM1	Q9rnm1 zymomonas m
742	6	2.2	78	10	Q9MSR3	Q9msr3 ipomoea bat	815	6	2.2	133	2	Q53795	Q53795 streptomyce
743	6	2.2	81	10	Q9MSR9	Q9msr9 ipomoea bat	816	6	2.2	133	2	Q93DR3	Q93dr3 trichodesmi
744	6	2.2	81	10	Q9MS55	Q9ms55 ipomoea bat	817	6	2.2	134	3	Q9URE2	Q9ure2 pneumocysti
745	6	2.2	81	16	Q8YEX1	Q8yex1 bruceella me	818	6	2.2	134	11	Q8C2Y7	Q8c2y7 mus musculu
746	6	2.2	83	8	Q37096	Q37096 hylobates l	819	6	2.2	134	13	Q57445	Q57445 oncorhynch

820	6	2.2	135	17	Q8TMF8	Q8TMF8 methanobarc	893	6	2.2	170	8	Q8SKH8	Q8SKH8 crocatus vi
821	6	2.2	136	10	Q94F68	Q94F68 tritricum ae	894	6	2.2	170	8	Q8SEI8	Q8SEI8 crocatus vi
822	6	2.2	136	16	Q9A1T7	Q9A1T7 caulobacter	895	6	2.2	170	8	Q8SKH9	Q8SKH9 crocatus vi
823	6	2.2	136	16	Q8F1V6	Q8F1V6 leptospira	896	6	2.2	170	8	Q8SKG7	Q8SKG7 crocatus vi
824	6	2.2	138	12	Q8J502	Q8J502 phthorimaea	897	6	2.2	170	8	Q8SKI5	Q8SKI5 crocatus vi
825	6	2.2	138	16	Q9RTC3	Q9RTC3 deinococcus	898	6	2.2	170	8	Q8SKH1	Q8SKH1 crocatus vi
826	6	2.2	139	17	Q59457	Q59457 pyrococcus	899	6	2.2	170	11	Q9DBF9	Q9DBF9 mus musculus
827	6	2.2	140	15	Q8SRV3	Q8SRV3 encephalito	900	6	2.2	171	5	Q9J3E3	Q9J3E3 caenorhabdi
828	6	2.2	140	16	Q8RCE1	Q8RCE1 thermotoga	901	6	2.2	171	16	Q8DNC6	Q8DNC6 streptococ
829	6	2.2	141	2	Q9J9J1	Q9J9J1 vibrio chol	902	6	2.2	172	2	Q9ZBB4	Q9ZBB4 streptococ
830	6	2.2	141	16	Q9WZ50	Q9WZ50 thermotoga	903	6	2.2	172	2	Q9J6H8	Q9J6H8 staphylococ
831	6	2.2	143	3	Q9FCC2	Q9FCC2 erwinia ste	904	6	2.2	172	16	Q8XRV2	Q8XRV2 ralstonia s
832	6	2.2	143	10	Q9FLB3	Q9FLB3 arabidopsis	905	6	2.2	172	16	Q93J00	Q93J00 streptomyce
833	6	2.2	143	17	Q9HP90	Q9HP90 halobacteri	906	6	2.2	172	16	Q8E7B9	Q8E7B9 streptococ
834	6	2.2	144	17	Q8Q0M0	Q8Q0M0 methanobarc	907	6	2.2	172	16	Q8E1Z1	Q8E1Z1 streptococ
835	6	2.2	145	11	Q91ZP7	Q91ZP7 rattus norv	908	6	2.2	173	12	Q8B6B1	Q8B6B1 aichi virus
836	6	2.2	145	16	Q8YAH9	Q8YAH9 listeria mo	909	6	2.2	173	16	Q24972	Q24972 helicobacte
837	6	2.2	147	15	Q8AU28	Q8AU28 human immu	910	6	2.2	173	17	Q27935	Q27935 methanobac
838	6	2.2	147	16	P96272	P96272 mycobacteri	911	6	2.2	174	11	Q8VC09	Q8VC09 mus musculi
839	6	2.2	148	17	Q8U2N2	Q8U2N2 pyrococcus	912	6	2.2	175	4	Q8N7F0	Q8N7F0 homo sapien
840	6	2.2	150	2	Q8Z099	Q8Z099 salmonella	913	6	2.2	177	16	Q9Z2E19	Q9Z2E19 listeria in
841	6	2.2	150	15	Q8S772	Q8S772 human immu	914	6	2.2	177	16	Q8Y994	Q8Y994 listeria mo
842	6	2.2	150	16	Q97NM4	Q97NM4 streptococ	915	6	2.2	178	2	Q9F4Z6	Q9F4Z6 shope fibro
843	6	2.2	150	16	Q8ZMY9	Q8ZMY9 salmonea	916	6	2.2	179	12	Q9Q8W3	Q9Q8W3 fusobacteri
844	6	2.2	151	16	Q8RAC2	Q8RAC2 thermobact	917	6	2.2	179	16	Q9ZVFI	Q9ZVFI arididopsis
845	6	2.2	152	2	Q46310	Q46310 carnobacter	918	6	2.2	180	10	Q9ZVFI	Q9ZVFI anabaena sp
846	6	2.2	152	17	Q9HIT9	Q9HIT9 thermopla	919	6	2.2	180	16	Q8XB64	Q8XB64 escherichia
847	6	2.2	153	16	Q8RF07	Q8RF07 fusobacteri	920	6	2.2	182	5	Q62209	Q62209 caenorhabdi
848	6	2.2	154	2	Q9ANN1	Q9ANN1 bradyrhizob	921	6	2.2	182	5	Q9FTW4	Q9FTW4 oryza sativ
849	6	2.2	154	16	P70835	P70835 borrelia bu	922	6	2.2	182	16	Q9WZ32	Q9WZ32 thermotoga
850	6	2.2	154	6	Q9X8K2	Q9X8K2 streptomyce	923	6	2.2	182	16	Q9PB31	Q9PB31 xylella fas
851	6	2.2	155	16	Q97931	Q97931 sus scrofa	924	6	2.2	184	5	Q22026	Q22026 caenorhabdi
852	6	2.2	155	10	Q9ZOF4	Q9ZOF4 arabidopsis	925	6	2.2	185	10	Q9ZQF2	Q9ZQF2 arabidopsis
853	6	2.2	156	2	Q44475	Q44475 azotobacter	926	6	2.2	185	10	Q8GT28	Q8GT28 lycopersico
854	6	2.2	156	16	P71780	P71780 mycobacteri	927	6	2.2	185	16	Q9ZMP9	Q9ZMP9 helicobacte
855	6	2.2	157	17	Q971W6	Q971W6 pseudomonas	928	6	2.2	186	16	Q9KND3	Q9KND3 vibrio chol
856	6	2.2	158	2	P72115	P72115 pseudomonas	929	6	2.2	186	16	Q9ABAO	Q9ABAO caulobacter
857	6	2.2	158	2	Q07102	Q07102 escherichia	930	6	2.2	186	16	Q8YLV6	Q8YLV6 anabaena sp
858	6	2.2	158	16	Q9PDJ9	Q9PDJ9 xylella fas	931	6	2.2	187	12	Q91LV2	Q91LV2 influenza a
859	6	2.2	159	16	Q9BFX3	Q9BFX3 mycoplasma	932	6	2.2	187	15	Q90CNA	Q90CNA human immu
860	6	2.2	159	16	Q8G7W6	Q8G7W6 bifidobacte	933	6	2.2	187	16	Q9BKJ7	Q9BKJ7 rhizobium l
861	6	2.2	159	16	Q8EN51	Q8EN51 oceanobacti	934	6	2.2	188	10	Q9X1J2	Q9X1J2 arabidopsis
862	6	2.2	160	2	Q9EY90	Q9EY90 corynebacte	935	6	2.2	188	16	Q8EDJ2	Q8EDJ2 shewanelia
863	6	2.2	160	2	Q9EUL3	Q9EUL3 corynebacte	936	6	2.2	188	16	Q8DCH5	Q8DCH5 vibrio vuln
864	6	2.2	160	2	Q8L368	Q8L368 salmonella	937	6	2.2	188	16	Q8DCH5	Q8DCH5 vibrio vuln
865	6	2.2	161	16	Q9H2J8	Q9H2J8 pseudomonas	938	6	2.2	189	11	Q8TWM2	Q8TWM2 methanopyru
866	6	2.2	162	1	Q50521	Q50521 methanobact	939	6	2.2	189	11	Q9CQH3	Q9CQH3 mus musculi
867	6	2.2	164	4	Q9NQ81	Q9NQ81 homo sapien	940	6	2.2	189	15	Q8BU20	Q8BU20 human immu
868	6	2.2	165	3	Q12157	Q12157 saccharomyc	941	6	2.2	189	15	Q9C9D6	Q9C9D6 arabidopsis
869	6	2.2	165	4	Q9BSC1	Q9BSC1 homo sapien	942	6	2.2	190	10	Q9C9D6	Q9C9D6 arabidopsis
870	6	2.2	165	4	Q8N9N8	Q8N9N8 homo sapien	943	6	2.2	190	16	Q97DH2	Q97DH2 clostridium
871	6	2.2	165	10	Q9FMC2	Q9FMC2 arabidopsis	944	6	2.2	191	5	Q8SBS5	Q8SBS5 encephalito
872	6	2.2	166	10	Q9LQJ9	Q9LQJ9 arabidopsis	945	6	2.2	191	15	Q904S7	Q904S7 human immu
873	6	2.2	166	10	Q9ZTE8	Q9ZTE8 arabidopsis	946	6	2.2	191	16	Q8XZ04	Q8XZ04 ralstonia s
874	6	2.2	166	10	Q9SBG5	Q9SBG5 arabidopsis	947	6	2.2	191	16	Q8VKL3	Q8VKL3 mycobacteri
875	6	2.2	166	16	Q9RXA7	Q9RXA7 deinococcus	948	6	2.2	191	16	Q8R838	Q8R838 thermobact
876	6	2.2	166	16	Q8F6N0	Q8F6N0 leptospira	949	6	2.2	192	10	Q8GTT1	Q8GTT1 symbiodinu
877	6	2.2	167	8	Q9MP7	Q9MP7 grallaria r	950	6	2.2	193	2	Q9JCV7	Q9JCV7 enterococu
878	6	2.2	167	8	Q9MP00	Q9MP00 grallaria s	951	6	2.2	193	6	Q28723	Q28723 oryctolagus
879	6	2.2	167	17	Q9MP8	Q9MP8 grallaria g	952	6	2.2	193	16	Q8YMN1	Q8YMN1 anabaena sp
880	6	2.2	168	16	Q9CEB8	Q9CEB8 lactococcus	953	6	2.2	194	2	Q9JEP7	Q9JEP7 photorhabdi
881	6	2.2	168	16	Q9CEB8	Q9CEB8 lactococcus	954	6	2.2	194	2	Q9JEP7	Q9JEP7 photorhabdi
882	6	2.2	168	16	Q8RB38	Q8RB38 thermobact	955	6	2.2	194	9	Q8WB67	Q8WB67 synchococ
883	6	2.2	168	16	Q8BPO0	Q8BPO0 corynebacte	956	6	2.2	194	9	Q8WB67	Q8WB67 synchococ
884	6	2.2	168	16	Q9YBR2	Q9YBR2 aeropyrum p	957	6	2.2	194	9	Q8WB67	Q8WB67 synchococ
885	6	2.2	169	11	Q64364	Q64364 mus musculi	958	6	2.2	194	16	Q904S9	Q904S9 human immu
886	6	2.2	170	8	Q8SKG9	Q8SKG9 crocatus vi	959	6	2.2	194	16	Q8NOR3	Q8NOR3 corynebacte
887	6	2.2	170	8	Q8SKH6	Q8SKH6 crocatus vi	960	6	2.2	196	2	Q9EYV7	Q9EYV7 rhizobium l
888	6	2.2	170	8	Q8SKI3	Q8SKI3 crocatus vi	961	6	2.2	196	16	Q9CHZ7	Q9CHZ7 lactococcus
889	6	2.2	170	8	Q8SKI1	Q8SKI1 crocatus vi	962	6	2.2	197	6	Q97673	Q97673 sus scrofa
890	6	2.2	170	8	Q8SKG5	Q8SKG5 crocatus vi	963	6	2.2	197	9	Q8SBY7	Q8SBY7 cyanophage
891	6	2.2	170	8	Q8SKH5	Q8SKH5 crocatus vi	964	6	2.2	197	9	Q8SBY6	Q8SBY6 cyanophage
892	6	2.2	170	8	Q8SKI4	Q8SKI4 crocatus vi	965	6	2.2	197	9	Q8SBY6	Q8SBY6 cyanophage



RC STRAIN=cv. V26; TISSUE=Petal limb;  
RX MEDLINE=99380006; PubMed=10449578;  
RA Quatrrochio F., Wing J., van der Woude K., Souer E., de Vetten N.,  
Mol J., Koes R.;  
RT "Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color.";  
RL Plant Cell 11:1433-1444(1999).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
DR EMBL: AF146702; AAF66727.1; -.  
DR HSSP: P06876; 1MBK.  
DR TRANSFAC: T02955; -.  
DR InterPro: IPR001005; Myb DNA binding.  
DR Pfam: PF00249; myb DNA-binding; 2.  
DR SMART: SM00717; SANT; 2.  
DR PROSITE: PS00037; MYB\_1; 1.  
DR PROSITE: PS00334; MYB\_2; 1.  
DR PROSITE: PS50090; MYB\_3; 2.  
KW DNA-binding; Nuclear protein.  
SQ SEQUENCE 255 AA; 29021 MW; 3C590E8473209A52 CRC64;  
  
Query Match 13.9%; Score 38; DB 10; Length 255;  
Best Local Similarity 100.0%; Pred. No. 8.8e-32;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 72 DEVDLIRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
72 DEVDLIRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
Db 72 DEVDLIRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
  
RESULT 4  
Q9M720 PRELIMINARY; PRT; 255 AA.  
AC Q9M720;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)  
DE AN2 protein.  
GN AN2.  
OS Petunia integrifolia (Violet-flowered petunia) (Petunia inflata).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; lamids; Solanales; Solanaceae; Petunia.  
OX NCBI\_TaxID=4103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. S9; TISSUE=Petal limb;  
RX MEDLINE=99380006; PubMed=10449578;  
RA Quatrrochio F., Wing J., van der Woude K., Souer E., de Vetten N.,  
Mol J., Koes R.;  
RT "Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color.";  
RL Plant Cell 11:1433-1444(1999).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
DR EMBL: AF146703; AAF66728.1; -.  
DR HSSP: P06876; 1MBK.  
DR InterPro: IPR001005; Myb DNA binding.  
DR Pfam: PF00249; myb DNA-binding; 2.  
DR SMART: SM00717; SANT; 2.  
DR PROSITE: PS00037; MYB\_1; 1.  
DR PROSITE: PS00334; MYB\_2; 1.  
DR PROSITE: PS50090; MYB\_3; 2.  
KW DNA-binding; Nuclear protein.  
SQ SEQUENCE 255 AA; 29007 MW; 48B18508A9D59C3F CRC64;  
  
Query Match 13.9%; Score 38; DB 10; Length 255;  
Best Local Similarity 100.0%; Pred. No. 8.8e-32;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 72 DEVDLIRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
72 DEVDLIRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
Db 72 DEVDLIRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109

RESULT 5  
Q9M716 PRELIMINARY; PRT; 127 AA.  
AC Q9M716;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)  
DE AN2 truncated protein.  
GN AN2.  
OS Petunia hybrida (Petunia).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; lamids; Solanales; Solanaceae; Petunia.  
OX NCBI\_TaxID=4102;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. W44; TISSUE=Petal limb;  
RX MEDLINE=99380006; PubMed=10449578;  
RA Quatrrochio F., Wing J., van der Woude K., Souer E., de Vetten N.,  
Mol J., Koes R.;  
RT "Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color.";  
RL Plant Cell 11:1433-1444(1999).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
DR EMBL: AF146707; AAF66732.1; -.  
DR HSSP: P06876; 1MBK.  
DR InterPro: IPR001005; Myb DNA binding.  
DR Pfam: PF00249; myb DNA-binding; 2.  
DR SMART: SM00717; SANT; 2.  
DR PROSITE: PS00037; MYB\_1; 1.  
DR PROSITE: PS00334; MYB\_2; 1.  
DR PROSITE: PS50090; MYB\_3; 2.  
KW DNA-binding; Nuclear protein.  
SQ SEQUENCE 127 AA; 14817 MW; F928569624274F35 CRC64;  
  
Query Match 12.4%; Score 34; DB 10; Length 127;  
Best Local Similarity 100.0%; Pred. No. 8.7e-28;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 76 LILRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
76 LILRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
Db 76 LILRLHKLGNRWSLIAGRLPGRTANDVKNYNT 109  
  
RESULT 6  
Q9M718 PRELIMINARY; PRT; 129 AA.  
AC Q9M718;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)  
DE AN2 truncated protein.  
GN AN2.  
OS Petunia hybrida (Petunia).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; lamids; Solanales; Solanaceae; Petunia.  
OX NCBI\_TaxID=4102;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. W15; TISSUE=Petal limb;  
RX MEDLINE=99380006; PubMed=10449578;  
RA Quatrrochio F., Wing J., van der Woude K., Souer E., de Vetten N.,  
Mol J., Koes R.;  
RT "Molecular analysis of the anthocyanin2 gene of petunia and its role  
in the evolution of flower color.";  
RL Plant Cell 11:1433-1444(1999).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
DR EMBL: AF146705; AAF66730.1; -.  
DR

DR HSSP; P06876; IMBK.  
 DR InterPro; IPR001005; Myb DNA binding.  
 DR Pfam; PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 DR DNA-binding; Nuclear protein.  
 DR SEQUENCE 129 AA; 15084 MW; 12D9DD38441D2427 CRC64;  
 SO

Query Match 12.4%; Score 34; DB 10; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-28;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 76 LRLHLKLLGNRWSLIAGRLPGRRTANDVKNYNT 109  
 76 LRLHLKLLGNRWSLIAGRLPGRRTANDVKNYNT 109

RESULT 7  
 Q9FNW8 PRELIMINARY; PRT; 139 AA.  
 AC Q9FNW8; 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Putative transcription factor MYB114 (Myb-related transcription factor, putative).  
 GN T27F4.13.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S., White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y., Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y., Gail J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A., Miltsecher J., Miranda M., Nguyen M., Nietman W.C., Osborne B.I., Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D., Utterback T., Van Aken S., Vaysberg M., Vyotskaya V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana."  
 RT Nature 408:816-820(2000).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
 DR EMBL; AY008379; AAG38381.1; -.  
 DR EMBL; AC020665; AAG52160.1; -.  
 DR HSSP; P06876; IIDY.  
 DR InterPro; IPR001005; Myb DNA binding.  
 DR Pfam; PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 DR SEQUENCE 129 AA; 15084 MW; 12D9DD38441D2427 CRC64;  
 SO

KW DNA-binding; Nuclear protein.  
 SQ SEQUENCE 139 AA; 16007 MW; 81D7B545F8274461 CRC64;  
 SO

Query Match 11.7%; Score 32; DB 10; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-25;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 78 LRLHLKLLGNRWSLIAGRLPGRRTANDVKNYNT 109  
 75 LRLHLKLLGNRWSLIAGRLPGRRTANDVKNYNT 106

RESULT 8  
 Q9FNW9 PRELIMINARY; PRT; 246 AA.  
 AC Q9FNW9; 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Putative transcription factor MYB113 (Myb-related transcription factor, putative).  
 GN T27F4.12.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S., White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y., Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y., Gail J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A., Miltsecher J., Miranda M., Nguyen M., Nietman W.C., Osborne B.I., Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D., Utterback T., Van Aken S., Vaysberg M., Vyotskaya V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana."  
 RT Nature 408:816-820(2000).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
 DR EMBL; AY008378; AAG38380.1; -.  
 DR EMBL; AC020665; AAG52158.1; -.  
 DR HSSP; P06876; IMBK.  
 DR InterPro; IPR001005; Myb DNA binding.  
 DR Pfam; PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 DR DNA-binding; Nuclear protein.  
 DR SEQUENCE 246 AA; 28307 MW; FF1C4B9176040D20 CRC64;  
 SO

Query Match 11.7%; Score 32; DB 10; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-25;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LRLHKLGNRWSLIAGRLPRTANDVKNYNT 109  
 Db 75 LRLHKLGNRWSLIAGRLPRTANDVKNYNT 106

## RESULT 9

Q92TC3 PRELIMINARY; PRT; 249 AA.  
 AC Q92TC3;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Putative transcription factor (Production of anthocyanin pigment 2 protein) (MYB-related transcription factor, putative, 65699-67047).  
 DE 67047).  
 GN MYB90 OR PAP2 OR T27F4.14.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eustoids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=99056849; PubMed=9839469;  
 RA Kranz H.D., Denekamp M., Greco R., Jin H.-L., Leyva A., Meissner R.,  
 RA Petroni K., Urzainqui A., Bevan M., Martin C., Smeekens S.,  
 RA Tonelli C., Paz-Ares J., Weisshaar B.,  
 RT "Towards functional characterisation of the members of the R2R3-MYB  
 RT gene family from Arabidopsis thaliana";  
 RL Plant J. 16:263-276(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia-0;  
 RX Borevitz J.O., Xia Y., Blount J., Dixon R.A., Lamb C.;  
 RT "Activation Tagging Identifies a Conserved MYB Regulator of  
 RT Phenylpropanoid Biosynthesis";  
 RL Plant Cell 12:0-0(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Lures J.S., Maizel R., Matzli A.,  
 RA Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Frazer C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 408:816-820(2000).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
 DR EMBL; AF062915; AAC83637.1; -;  
 DR EMBL; AF325124; AAG42002.1; -;  
 DR EMBL; AC020665; AAG52164.1; -;  
 DR HSSP; P06876; 1IDY.  
 DR InterPro: IPR001005; Myb DNA binding.  
 DR Pfam: PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR SMART; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.

KW DNA-binding; Nuclear protein.  
 SQ SEQUENCE 249 AA; 28100 MW; 309024FD4103700A CRC64;

Query Match 11.7%; Score 32; DB 10; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-25;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LRLHKLGNRWSLIAGRLPRTANDVKNYNT 109  
 Db 75 LRLHKLGNRWSLIAGRLPRTANDVKNYNT 106

## RESULT 10

Q9M719 PRELIMINARY; PRT; 255 AA.  
 AC Q9M719;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE AN2 protein.  
 GN AN2.  
 OS Petunia integrifolia (Violet-flowered petunia) (Petunia inflata).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamids; Solanales; Solanaceae; Petunia.  
 OC NCBI\_Taxid=4103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. S6; TISSUE=Petal limb;  
 RX MEDLINE=9380006; PubMed=10449578;  
 RA Quattrocchio F., Wang J., van der Woude K., Souer E., de Vetten N.,  
 RA Mol J., Koes R.;  
 RT "Molecular analysis of the anthocyanin2 gene of petunia and its role  
 RT in the evolution of flower color";  
 RL Plant Cell 11:1433-1444(1999).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
 DR EMBL; AF146704; AAF66729.1; -;  
 DR HSSP; P06876; 1MBX.  
 DR InterPro: IPR001005; Myb DNA binding.  
 DR Pfam: PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KW DNA-binding; Nuclear protein.  
 SQ SEQUENCE 255 AA; 28963 MW; E4626F08B07071AD CRC64;

Query Match 10.2%; Score 28; DB 10; Length 255;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-21;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 RAGLNCRKSCRRLMYLRLPHIKRGDF 69  
 Db 42 RAGLNCRKSCRRLMYLRLPHIKRGDF 69

## RESULT 11

Q92TC7 PRELIMINARY; PRT; 212 AA.  
 AC Q92TC7;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Putative transcription factor.  
 GN MYB75.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eustoids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;



RA MEDLINE=99056848; PubMed=9839469;  
RA Kranz H.D., Denekamp M., Greco R., Jin H., Leyva A., Meissner R.C.,  
RA Petrini K., Uzzauini A., Bevan M., Martin C., Smeekens S.,  
RA Tonelli C., Paz-Ares J., Weishaar B.,  
RT "Towards functional characterisation of the members of the R2R3-MYB  
RT gene family from Arabidopsis thaliana.",  
RL Plant J. 16:263-276(1998).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC EMBL: AF062908; AAC83630.1; -.  
DR HSPSP; P06876; 11DY.  
DR InterPro: IPR001005; MYB\_DNA\_binding.  
DR Pfam: PF00249; myb\_DNA-binding; 2.  
DR SMART: SM00717; SANT; 2.  
DR PROSITE: PS00037; MYB\_1; 1.  
DR PROSITE: PS00334; MYB\_2; 1.  
DR PROSITE: PS50090; MYB\_3; 2.  
KW DNA-binding; Nuclear protein.  
SQ SEQUENCE 212 AA; 24508 MW; DE53594D17AA02A0 CRC64;  
  
Query Match 9.9%; Score 27; DB 10; Length 212;  
Best Local Similarity 100.0%; Pred. No. 3.9e-20;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 83 LIGNRWSLIAGRLPGRTANDVKYMYNT 109  
Db 80 LIGNRWSLIAGRLPGRTANDVKYMYNT 106  
|||||  
AC Q9FE25 PRELIMINARY; PRT; 248 AA.  
ID Q9FE25  
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)  
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)  
DT 01-MAR-2003 (TREMblrel. 23, last annotation update)  
DE Production of anthocyanin pigment 1 protein (Putative transcription  
DE factor).  
GN PAP1 OR P25P12.92.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia-0;  
RA Borevitz J.O., Xia Y., Blount J., Dixon R.A., Lamb C.;  
RT "Activation Tagging Identifies a Conserved MYB Regulator of  
RT Phenylpropanoid Biosynthesis.",  
RL Plant Cell 12:0-0(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Federlepie N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,  
RA Alalati H., Nguyen M., Lam B., Southwick A., Bel, Buehler E., Chin C.,  
RA Chou J., Choi E., Dunn P., Gonzalez A., Hwang B., Kim C., Koo T.,  
RA Lee J.M., Lenz C., Li J., Liu A., Liu K., Liu S., Mubharaky N.,  
RA Pham P., Sakano H., Schwartz J., Shinn P., Thaverei A., Toriumi M.,  
RA Vayberg M., Walker M., Yu G., Becker J., Theologis A., Davis R.W.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC EMBL: AF135123; AAG42001.1; -.  
DR EMBL: AC009323; AAG09100.1; -.  
DR HSPSP; P06876; 11DY.  
DR InterPro: IPR001005; MYB\_DNA\_binding.  
DR Pfam: PF00249; myb\_DNA-binding; 2.  
DR SMART: SM00717; SANT; 2.  
DR PROSITE: PS00037; MYB\_1; 1.  
DR PROSITE: PS00334; MYB\_2; 1.  
DR PROSITE: PS50090; MYB\_3; 2.  
KW DNA-binding; Nuclear protein.  
SQ SEQUENCE 248 AA; 28469 MW; B86103A9AB5CF442 CRC64;

Query Match 9.9%; Score 27; DB 10; Length 248;  
Best Local Similarity 100.0%; Pred. No. 4.5e-20;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 83 LIGNRWSLIAGRLPGRTANDVKYMYNT 109  
Db 80 LIGNRWSLIAGRLPGRTANDVKYMYNT 106  
|||||  
AC Q94FG4 PRELIMINARY; PRT; 179 AA.  
ID Q94FG4  
DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)  
DT 01-MAR-2003 (TREMblrel. 23, last annotation update)  
DE MYB-like transcription factor Myb 3 (Fragment).  
OS Gossypium herbaceum.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
OX NCBI\_TaxID=34274;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cron R., Cedroni M., Haselkorn T., Osborne C., Wendel J.F.;  
RT "PCR-mediated recombination in a polyploid plant.",  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC EMBL: AF377306; AAK58019.1; -.  
DR InterPro: IPR001005; MYB\_DNA\_binding.  
DR Pfam: PF00249; myb\_DNA-binding; 2.  
DR SMART: SM00717; SANT; 2.  
DR PROSITE: PS00037; MYB\_1; 1.  
DR PROSITE: PS00334; MYB\_2; 1.  
DR PROSITE: PS50090; MYB\_3; 2.  
KW Nuclear protein.  
FT NON TER 1  
SQ SEQUENCE 179 AA; 20830 MW; CAA70CA27FC9A8F CRC64;  
  
Query Match 9.1%; Score 25; DB 10; Length 179;  
Best Local Similarity 100.0%; Pred. No. 4.5e-18;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 75 DLILRLKLIAGRLPGRT 99  
Db 62 DLILRLKLIAGRLPGRT 86  
|||||  
AC Q8H262 PRELIMINARY; PRT; 181 AA.  
ID Q8H262  
DT 01-MAR-2003 (TREMblrel. 23, Created)  
DT 01-MAR-2003 (TREMblrel. 23, last sequence update)  
DT 01-MAR-2003 (TREMblrel. 23, last annotation update)  
DE MYB-like transcription factor 3 (Fragment).  
GN MYB3.  
OS Gossypium kirkii.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Malvales; Malvaceae; Malvoideae; Gossypioideae.  
OX NCBI\_TaxID=47615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cedroni M.L., Cron R.C., Adams K.L., Wilkins T.A., Wendel J.F.;  
RT "Evolution and expression of MYB genes in diploid and polyploid  
RT cotton".  
RL Plant Mol. Biol. 0:0-0(2002).  
DR EMBL: AY115511; AAN28279.1; -.  
FT NON TER 1  
SQ SEQUENCE 181 AA; 20967 MW; 0F0A0FDE13250F7D CRC64;



Query Match 9.1%; Score 25; DB 10; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-18;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 DILRLHKLLGNRWSLIAGRLPGRT 99  
 |||||  
 DB 64 DILRLHKLLGNRWSLIAGRLPGRT 88

## RESULT 15

Q94FG5 PRELIMINARY; PRT; 184 AA.  
 AC Q94FG5;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)  
 DE Myb-like transcription factor Myb 3 (Fragment).  
 OS Gossypium raimondii.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
 OX NCBI\_TaxID=29730;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cron R., Cedroni M., Haselkorn T., Osborne C., Wendel J.F.;  
 RT "PCR-mediated recombination in a polyploid plant."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.  
 DR EMBL; AF377305; AAK58018.1; -;  
 DR InterPro; IPR001005; MYB DNA binding.  
 DR Pfam; PF00249; myb DNA-binding; 2.  
 DR SMART; SM00717; SMT; 2.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KW Nuclear protein.  
 FT NON TER 1  
 SQ SEQUENCE 184 AA; 21435 MW; FEEDB20F1E3E407A CRC64;

Query Match 9.1%; Score 25; DB 10; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-18;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 DILRLHKLLGNRWSLIAGRLPGRT 99  
 |||||  
 DB 67 DILRLHKLLGNRWSLIAGRLPGRT 91

Search completed: January 29, 2004, 20:21:09  
 Job time : 70 secs

